

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 2, 2003, 04:55:28 ; Search time 272 Seconds
(without alignments)
822.725 Million cell updates/sec

Title: US-09-977-260-2
Perfect score: 2671
Sequence: 1 MACRGLSVWRAFHGCDSEAE.....PASVSGODADGSTRSPRQEP 507

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n_model -DEV=xip
-O=/cgn2_1/USPTO.spool/US09977260/runtat_29072003_092741_14649/app_query.fasta.1.647
-DB=issued.patents_NA -Qfmt=fastap -SUFFIX=ini -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALL=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US09977260.ecgn_1.1.85.gunat_29072003.092741_14649 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NBS_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents.NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2671	100.0	1942	2	US-08-604-989A-11
2	2671	100.0	2000	4	US-08-426-509A-1
3	2671	100.0	2000	4	US-08-232-545-1
4	2671	100.0	2000	5	PCT-US93-05008-1
5	2662	99.7	1521	2	US-08-604-989A-10
6	2591.5	97.0	1987	2	US-08-876-882-1
7	2591.5	97.0	1987	4	US-09-315-928-1
8	2444	91.5	1398	2	US-08-604-989A-9
9	2020.5	75.6	1713	4	US-09-741-154-1
10	1469.5	55.0	16389	2	US-08-741-154-3
11	1269	47.5	738	2	US-08-604-989A-8
12	737.5	27.6	1602	1	US-07-820-011A-1

13	737.5	27.6	1602	5	PCT-US93-00445-1
14	727	27.2	1611	1	US-07-820-011A-3
15	727	27.2	1611	5	PCT-US93-00445-3
16	727	27.2	2129	4	US-09-016-434-1452
17	710	26.6	4517	5	PCT-US93-06251-83
18	702.5	26.3	2647	4	US-09-220-132-77
19	702.5	26.3	2647	5	PCT-US93-06251-77
20	691.5	25.9	1491	2	US-09-006-675-1
21	691.5	25.9	1491	3	US-09-228-603A-1
22	683.5	25.6	3633	1	US-08-306-691B-35
23	680	25.5	1467	4	US-09-579-182-2
24	680	25.5	1548	4	US-09-099-053-1
25	656	24.6	2827	4	US-08-492-723-1
26	654.5	24.5	2770	4	US-08-426-509A-5
27	654.5	24.5	2770	5	US-08-232-545-5
28	654.5	24.5	2770	5	PCT-US95-05008-5
29	654.5	24.5	7607	1	US-08-222-616-19
30	654.5	24.5	7607	4	US-08-446-648-19
31	654.5	24.5	7607	5	PCT-US95-04228-19
32	608.5	22.8	2674	4	US-09-817-180-1
33	607	22.7	2505	1	US-08-391-615-1
34	595.5	22.3	3503	1	US-07-631-717A-1
35	595.5	22.3	3503	1	US-08-166-717D-1
36	591	22.1	1875	1	US-08-070-165F-3
37	591	22.1	1875	2	US-08-885-418-3
38	587	22.0	2456	4	US-09-016-434-1476
39	587	22.0	2500	4	US-08-426-509A-3
40	587	22.0	2500	4	US-08-232-545-3
41	587	22.0	2500	5	PCT-US95-05008-3
42	582	21.8	2049	4	US-09-099-749-10
43	581.5	21.8	5993	3	US-09-383-630-1
44	581.5	21.8	5993	3	US-09-383-630-2
45	576.5	21.6	2469	1	US-07-997-133-2

ALIGNMENTS

RESULT 1
US-08-604-989A-11
Sequence 11, Application US/08604989A
Patent No. 5834208

GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: NO. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Charles F. Miller

REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 1942 base pairs
TYPE: nucleic acid

```

STRANDEDNESS: double
;
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN: UT-7
US-08-604-989A-11

Alignment Scores:
Pred. No.: 2,17e-257 Length: 1942
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-977-260-2 (1-507) x US-08-604-989A-11 (1-1942)
QY 1 MetAlaGlyArgGlySerLeuValSerTTPArgAlaPheHisGlyCysAspSerAlaGlu 20
Db 208 ATGGCGGGGAGAGGCTCTGTTCCTGGCGGGCAATTCACGCGTGTCTGCTGAG 267
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTPrHisProProValSerAla 40
Db 268 GAACCTCCCGGGGTGAGCCCGCTTCCTCCGAGCTGGCACCCTCCCGCTCAGCC 327
QY 41 ArgMetProThrArgArgTPrAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
Db 328 AGAGTCCCAACGAGCGCTGGGCGCCGCGGACCCAGTGTATCCAAATGCGAGCACACC 387
QY 61 ArgProLysProGlyLysLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
Db 388 CGGCCCAAGCGAGGGAGCTGGCTTCGCCAAGGCGAGGTGTGTCACCATCTGGAGGCC 447
QY 81 CysGluAsnLysSerTPrArgValLysHisHisThrSerGlyGlnGlnGlyLeuLeu 100
Db 448 TGGGACACAAAGAGCTGTACCGCCGTCACAGCACACACAGTGTGAGAGAGGGCTGCTG 507
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 508 GAGGCTGGGGCGGTGGGAGCGGAGCGGCTCTCCGACAGCCCAAGCTCAGCCCTCAGT 567
QY 121 ProTPrPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
Db 568 CGGTGGTTCACGCGAGAGATCTCGGGCCAGAGGCTGTCCAGAGCTGACGCTCCCGAG 627
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysVal 160
Db 628 GATGGCTGTCTCTGGTGGGAGAGTCCGCGCGCCACCCGCGGAGCTACGCTCTGCTG 687
QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
Db 688 AGCTTTGGCGCGGACGTCACTCCAGCTCCGCTGCACCGGAGCGCCACCTCACAATC 747
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
Db 748 GATGAGCGCGGTCTCTGTGCAACCTCAGAGCATGTGTGAGATTAACAGCAAGCAAG 807
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
Db 808 GCGCGATCTGCACCAAGCTGTGTGAGACCAAAAGCGGAAACACGCGCAAGTGGCGGAG 867
QY 221 GlnGluLeuAlaArgAlaGlyTPrLeuLeuAsnLeuGlnHisLeuThrLeuGlnAlaGln 240
Db 868 GAGGAGCTGGCCAGCGCGGTGTGTAGAACCTGCACATTTGACATTTGGAGACAG 927
QY 241 IleGlyGlnGlyGlnPheGlyAlaValIleGlnGlnGlyLysIleGlnLysValAla 260
Db 928 ATCGGAGAGGAGAGTGTGAGCTGTCTGCAAGGGGTGAGTACCTGGGCAAAAGGTGGCC 987
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
Db 988 GTGAAGATATTCACAGTGTATGTGACAGCCAGCGCTTCTGTGACAGACGCGCGCTCAG 1047

```

```

QY 281 ThrLysMetGlnHisGlnAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlnLeu 300
Db 1048 ACAGAGATGACAAACAGAGAACTGTGTGCTCTCTGGGGTGTATCTCTGCACACAGGGGCTG 1107
QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgTPrArgGly 320
Db 1108 TACATTGTGATGAGACAGCTGTGAGCAAGGGCCAAAGCTGTGTGAGTGTGTGGGAGCCGGGGT 1167
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMet 340
Db 1168 CGAGCCCTGTGTACACACCGCTCAGCTCTCTGCAATTTCTCTGACAGCTGGCCGAGGCAAG 1227
QY 341 GluTyrLeuGlnSerLysLysLeuValHisArgAspLeuAlaIleArgAsnIleLeuVal 360
Db 1228 GAGTACCTGGAGAGCAAGAACTTGTGACCCGCGAGCTGGCCCGCCGCAACATCTGTGTC 1287
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
Db 1288 TCAGAGGACCTGTGTGCGCAAGGTCAAGCTGTGTGGCTGGCCAAAGCCGAGGAGGAG 1347
QY 381 LeuAspSerSerArgLeuProValLysTPrThrAlaProGluAlaLeuLysHisGlyLys 400
Db 1348 CTAGACTCAAGCCGCGCTGCCCTCAGTGTGAGCGCGCCGAGGCTCTCAACACGGGAG 1407
QY 401 PheThrSerLysSerAspValTPrSerPheGlyValLeuLeuTPrGluValPheSerTyr 420
Db 1408 TTCACACGACAAAGCTGATGTCTGAGTTTGGGGTCTCTCTGGGAGGCTCTTCATAT 1467
QY 421 GlyArgAlaProThrProLysMetSerLysGlnValSerGluAlaValGlnGlyLysGly 440
Db 1468 GAGCGGGCTCCGTAACCTTAATGTACATGAAGAAGTGTGAGAGCCGTGGAGAAAGGG 1527
QY 441 TyrArgMetGlnProProGlnGlyCysProGlyProValHisValLeuMetSerSerCys 460
Db 1528 TACCGATGGAACCCCGCGAGGCTGTCCAGGCGCGCTGTCTCATAGACACTGCG 1587
QY 461 TPrGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db 1588 TGGGAGGAGAGAGCCCGCCGCGCCAGCCCTTCGCAAACTGGCCAGAGAGCTGGCCCGG 1647
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
Db 1648 GAGCTAGGAGTGTGAGTGTGCGCCAGCTGTGTGAGGAGGAGCGCCGAGCTTCACCC 1707
QY 501 SerProArgSerGlnGluPro 507
Db 1708 TCGCCCGCAAGCCAGAGAGCC 1728

RESULT 2
US-08-426-509A-1
; Sequence 1, Application US/08426509A
; Patent No. 6326469
; GENERAL INFORMATION:
; APPLICANT: Ulirich, Axel
; APPLICANT: Gishizky, Mikhail
; APPLICANT: Sures, Irman G.
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
; TITLE OF INVENTION: TYROSINE KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York,
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/426,509A
FILING DATE: 21-APR-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/232,545
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: COITZEL, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-0074-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNTE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-426-509A-1

Alignment Scores:
Pred. No.: 2,27e-257 Length: 2000
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-977-260-2 (1-507) x US-08-426-509A-1 (1-2000)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
DB 258 ATGGGGGGGGCGAGGCTCTGCTGCTGCGGCGCATTTTCACGGCTGATTCGCTGAG 317
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
DB 318 GAACTTCCCGGGGTAGCCCGGCTCTCCGAGGCTGAGCCCGCCCTCCGCTCACCC 377
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrIleCysGlnHisThr 60
DB 378 AGGATGCCACAGAGCGCTGGGCGCCCGGCGACCCAGTATCACCACCAATGGAGACACC 437
QY 61 ArgProLysProGlyGluLeuAlaPheArgIleGlyAspValValThrIleLeuGluAla 80
DB 438 CGCCCAAGCCAGGGAGCTGGCCCTCCGCAAGGCGAGCTGTCACCATCCTGGAGGCC 497
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGlyLeuLeu 100
DB 498 TGGCGAGACAGAGCTGATCCGCTCAAGCACACACAGAGGGAGGGGCTGCTG 557
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
DB 558 GCACCTGGGGCGCTGGCGGAGCGGAGGCGCTCTCCGAGACCCCAAGCTCAGCCTCATG 617
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
DB 618 CCGTGGTTCACGGGAAATCTCGGGCAGAGGCTGTCACGACGTGACACCTCCCGAG 677
QY 141 AspGlyLeuPheLeuValArgLysSerAlaArgHisProGlyAspTrpValLeuCysVal 160
DB 678 GATGGGCTGTTCGTGGTGGAGATCCGCGCGGCGCACCCGGCGACTGATCTGTGCTG 737
QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
DB 738 AGCTTTGGCCGCGAGCATCATCCATCCGCTGTCAGCCGCGAGCCACCTCAACATC 797
QY 181 AspGluAlaValIlePhePheCysAsnLeuMetAspMetValGlnHisTrpSerLysAspLys 200
DB 798 GATAGGCGCGTCTTCTGCAACCTATGACATGTTGGAGCAATTCACAGCAAGACACAG 857
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220

DB 858 GGGCGTATGTCACCAAGCTGGTGAAGACCAAAAGCGGAAACAGGACCAAGTCGGCGAG 917
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGluAlaGln 240
DB 918 GAGGAGCTGGCCAGGGCGGGTGGTTACTGAACCTGACACATTTGATTTGGAGACAG 977
QY 241 IleGlyGlnGlyLysPheGlyAlaValLeuGlnGlnGlyTrpLeuGlyGlnLysValAla 260
DB 978 ATCGGAGAGGAGAGATTTGGAGCTGCTCGCAGGCTGAGTACCTGGGCGAAAGGTGGCC 1037
QY 261 ValIysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
DB 1038 GTGAAGAAATATCAATGTGATGTGACAGCCAGGCTCTCTGAGAGAGGCGCTCATG 1097
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
DB 1098 ACGAAGATGCCAACAGAGAACCTGGTGGCTCTGCGCGTATCTGCACACAGGGGCTG 1157
QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
DB 1158 TACATTGTATGAGACAGCTGAGCAAGGCAACCTGTGAACTTTCTCGGACCGGGGT 1217
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluIleMet 340
DB 1218 CGAGCCCTCGTGAACACACCGCTCAGCTCTGAGTTTCTGTGACAGTGGCCGAGGGCATG 1277
QY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaIleAsnIleLeuVal 360
DB 1278 GAGTACCTGGAGAGCAAAACCTGTGACCGCGCGCTGCGCGCCCGCAACATCTCGTGC 1337
QY 361 SerGluAspLeuValAlaValSerAspPheGlyLeuAlaLysAlaGluArgGly 380
DB 1338 TCAGAGACATCGTGTCGCAAGTCAAGCATTTGGCTGGCCAAAGCCGAGCGGAGGGG 1397
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB 1398 CTAGACTCAACCGGCTCGCTCCGTCAGAGGAGCGCGGAGGCTCTCAAAACAGGGAG 1457
QY 401 PheTrpSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420
DB 1458 TTCACACAGAGTGGAGTGTGGAGTTTGGGGTGGCTGCTGGAGGCTTCTCAAT 1517
QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGluValSerGluAlaValGluGly 440
DB 1518 GGAAGGGCTCGTACCTCAAAATGTCACCTGAAAGAGTGTGGAGCGGTGGAAGAAGGG 1577
QY 441 TyrArgMetGluProProGlnGlyCysProGlyProValHisValLeuMetSerCys 460
DB 1578 TACCGCATGAAACCCCGGAGGCTGTCCAGGCGCCGTCACGTCTCATGAGCAGCTGC 1637
QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
DB 1638 TGGGAGCAGAGCCCGCCCGCCGACCTTCGCAAACTGGCGGAGAAAGCTGGCCGG 1697
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
DB 1698 GAGCTAGCAGTGTGAGGTGCCACAGCTCGTCTCAGGGCAGAGCGCGAGGCTCCAC 1757
QY 501 SerProArgSerGlnGluPro 507
DB 1758 TCGCCCGAAGCCAGAGGCC 1778

RESULT 3
US-08-232-545-1
Sequence 1, Application US/08232545
Patent No. 6506578
GENERAL INFORMATION:
APPLICANT: Ullrich, Axel
APPLICANT: Gishizsky, Mikhail
APPLICANT: Sures, Irmann G.
TITLE OF INVENTION: No. 6506578el Megakaryocytic Protein Tyrosine
TITLE OF INVENTION: Kinases

NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,545
FILING DATE: 22-Apr-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-232-545-1

Alignment Scores:
Pred. No.: 2,27e-257 Length: 2000
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-977-260-2 (1-507) x US-08-232-545-1 (1-2000)

QY 1 MetalGlyArgGlySerLeuValSerTPArgAlaPheHisGlyCysAspSeraIaGlu 20
DB 258 ATGGCGGGCGAGGCTCTCTGGTTTCCTGGCGGCGCATTTACAGCGTGATTCGTGAG 317
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSeraIa 40
DB 318 GAACCTTCCCGGGGTGAGCCCGCTTCCTCCGAGCTGGCACCCCTCCCGTCTCAGCC 377
QY 41 ArgMetProThrArgTrpAlaProGlyThrGlnCysIleThrIysCysGluHisThr 60
DB 378 AGGATCCCAAGAGGGGCTGGGCCCGGCGCACCATGATATCCAAATGCGAGCACACC 437
QY 61 ArgProLysProGlyLysLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
DB 438 CGCCCCAAGCCAGGGGAGCTGGCTCCGCAAGGGCGAGTGTGTCACCATCTCGAGGCC 497
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluLeuLeu 100
DB 498 TGGGAGACAAAGAGCTGTACCGCTCAAGCACACACAGTGGACAGAGGGGCTGTG 557
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
DB 558 GCGAGCTGGGGCGGTGGGGAGCGGAGGCCCTTCGCGAGACCCCAAGCTCAGCTCATG 617
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
DB 618 CCGTGTGTTCCAGCGGAAATCTCGGGCCAGAGGCTGTCCAGAGCTGCAGGCTCCGAG 677
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160

DB 678 GATGGGCTGTCTCGGCGGAGCTCGCGGCCACCCCGGCGACTACGCTGTGCTG 737
QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
DB 738 AGCTTTGGCGCGCACGTCATCCACTACCGGGTCTGACCGGCGACCGCCACTCATATC 797
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTrpSerLysAspLys 200
DB 798 GATGAGGCGCGTGTCTTCTCAACCTCATGATGATGGTGGAGCATTAACAGAGCAAG 857
QY 201 GlyAlaIleCysThrIysLeuValArgProLysArgLysHisGlyThrIysSeraIaGlu 220
DB 858 GCGGCTATCTGCACCAAGCTGTGAGACCAACCGGAAACCGGAGCAAGTGGCGGAG 917
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
DB 918 GAGGAGCTGGCCAGGGCGGGCTGTACTACACTGACATTTGACATTTGGAGCACAG 977
QY 241 IleGlyGluGlyLysPheGlyAlaValLeuGlnGlyLysIleGlyLeuGlyValAla 260
DB 978 ATCGAGAGAGGAGATTGGAGCTGTCTGACAGGTGAGTACCTGGGGCAAAAGTGGCC 1037
QY 261 ValLysAsnLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
DB 1038 GTGAGAAATATCAAGTGTATGTGACAGCCGCTTCTGTGACAGAGAGCGGCTCATG 1097
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
DB 1098 ACGAAGATGCACACACAGAACTGTGCTGCTCTG6GCGTATCTTCGACACAGGGGCTG 1157
QY 301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgLys 320
DB 1158 TACATTTGTATGAGACAGACGAGCAAGGCAACCTGTGAATTTCTGGGAGACCGGGGT 1217
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
DB 1218 CGAGCCCTGTGAAACACGCTGAGCTTCGCAATTTTCTGTGACCTGGCCGCGGCGATG 1277
QY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaIlaArgAsnIleLeuVal 360
DB 1278 GAGTACTGAGAGCAAGAACTGTGTCACCGGAGCTGCGCCCGGCAACATCTCTGCTC 1337
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
DB 1338 TCAGAGGAGCTGTGTGGCCAAAGTCAAGCTTGTGGCTGCGCAAAACCGAGGGAAGGG 1397
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB 1398 CTAGACTCAAGCGGCTGCCCGTCAAGTGTGAGCGCGCCGAGGCTCTCAAAACCGGGAAG 1457
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
DB 1458 TTCACACAGCAAGTCGATGTCTGAGTTTGGGGTCTCTCTGSGAGGTCTTCTCATAT 1517
QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGlyValSerGluAlaValGluLysGly 440
DB 1518 GGAGGGGCTCCGATCCCTAAATGTCACTGAAGAGAGTGTGGAGGCGCTGAGAAAGGGG 1577
QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerCys 460
DB 1578 TACCGCATGAAACCCCGGAGGGCTGTCCAGGCGCCCTGACGTCTCTTGTGAGCAAGTTC 1637
QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
DB 1638 TGGGAGGAGAGACCCCGCGGCGGACACCTTCGCAAACTGGCCAGAGAGTGGCGCGG 1697
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspLysSerThr 500
DB 1698 GAGCTTAGCGAGTACAGGTCCCAAGCTCGTCTGAGGAGAGAGAGCGGAGGCTTCACAC 1757
QY 501 SerProArgSerGlnPro 507
|||||

Db 1758 TCGCCCCGAAAGCCAGAGCC 1778

RESULT 4
PCT-US95-05008-1
Sequence 1, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Missenschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
NUMBER OF INVENTIONS: 21
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNITE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
PCT-US95-05008-1

Alignment Scores:
Pred. No.: 2 276-257 Length: 2000
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-977-260-2 (1-507) x PCT-US95-05008-1 (1-2000)

QY 1 MetalaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
Db 258 ATGGGGGGGCGAGGCTCTCTGTTCCCTGGCGGCGCATTTTCACGGCTGTGATTTCTGCTAG 317
QY 21 GluLeuProAlaGlySerProAlaGlyPheLeuAlaGlyAlaTrpHisProProValSerAla 40
Db 318 GAACCTCCCGGGGTAGCCCGCGCTTCTCCGAGCGCTGCACCCCGCTCCGCTCACCC 377
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60

Db 378 AGGATGCCAACGAGGGCTGGGCCCCGGGGCACCACCTGATATCCAAATGCCAGCACACC 437
QY 61 ArgProLysProGlyGluLeuAlaPheArgGlyGlyAspValAlaThrIleLeuGluAla 80
Db 438 CGCCCCAACGACGAGGGAGCTGGCTTCCGCAAGGGCGAGCTGTCACATCTGGAGGGCC 497
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisIleThrSerGlyGlnGluLeuLeu 100
Db 498 TCGGAGAACAGAGCTGGTACCGCGTCAAGCACACCCAGTGAGCAGAGGGGCTGTG 557
QY 101 AlaAlaGlyAlaLeuArgGluAlaArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 558 GCAGCTGGGGCGCTGGGGAGGGAGGGCGCCCTCCGCGACCCCAAGCTCAGCTCAG 617
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaGlnGluLeuInProGlu 140
Db 618 CCGTGGTTCACGAGGAGATCTCGGGCCAGGAGGCTGTCCACACCTGCAGCTCCCGAG 677
QY 141 AspGlyLeuPheLeuValArgLysSerAlaArgHisProGlyAspTrpValLeuGlyVal 160
Db 678 GATGGGCTTCTCTGGTGGGAGTCCCGCGCCACCCCGGACTACGTCCTGTGCTG 737
QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisIleThrIle 180
Db 738 ACCTTGGCGCGGAGCTGATCCAGCGCTGTGACACCGCGAGGGCCACTCACAATC 797
QY 181 AspGluAlaValaPhePheCysAsnLeuMetAspMetValGlnHisTrpSerLysAspLys 200
Db 798 GATGAGGCCGCTTCTTCTGCAACCTCATGACATGAGTGGACATTCAGCAGAGGACAAG 857
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
Db 858 GCGGCTATCTGCACCAAGCTGTGAGACCAAGCGGAAACCGGAGCAAGTCCGCCGAG 917
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisIleThrLeuGlyAlaGln 240
Db 918 GAGGAGCTGGCGAGGGCGGCTGTACTGAGCACTGCAGCATTTGACATTTGGGACACAG 977
QY 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlnGlyTrpLeuGlyGlnLysValAla 260
Db 978 ATCGGAGAGGAGAGATTTGGAGCTGTCTGCGAGGCTGTGATACCTTGGGCAAAAGTGGCC 1037
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
Db 1038 GTGAAGATATTAATGATGATGTGACACCCAGGCTTCTGTGAGAGAGCGGCTCATG 1097
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
Db 1098 ACGAAGATGCACACAGAACACTGTGGCTGCTCGGCGGTGATCTGCACAGGGGCTG 1157
QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
Db 1158 TACATTGTCAATGAGACAGTGTGACAGGCAACCTGTGTAACTTTCTCGGAGGGGT 1217
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
Db 1218 CGAGGCGCTGTGAAACACCGGCTCAGCTGTGAGTTTCTGTGACGTGGCCGAGGGCATG 1277
QY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
Db 1278 GAGTACCTGTGAGAGCAAGAACCTTGTGCACCGGCACTGTGGCGGCGGCAACATCTGTGTC 1337
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
Db 1338 TCAAGAGACTGTGTGGCCAAAGGTCAAGCATTTGTGGCTGTGGCAAGCGGAGGAGGG 1397
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
Db 1398 CTAGACTCAACCGGCTCCGCTCAAGTGGAGCGGCGCCGAGGCTCTAAACACAGGGAG 1457
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420

Db 1458 TTCACGACGAGTGGATGTCTGGAGTTTGGGCTGCTCTCTGGAGCTTCTCATAT 1517
Qy 421 GYARGLAARProTyProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
Db 1518 GGAGGGGCTCGCTACCTAAATGTCACTGAAGAGGTGTGGAGGGCGCTGGAGAGGGG 1577
Qy 441 TyTrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerCys 460
Db 1578 TACCGCATGGAACCCCGAGAGGCTGTCCAGGCCCGCTGCACGTCTCATGAGCAGCTGC 1637
Qy 461 TTPGLUAGLUProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db 1638 TGGGAGGACAGAGCCCGCCCGCCGACCTTCCGCAACTGGCCGAGAGAGTGGCCCGG 1697
Qy 481 GluLeuArgSerAlaGluAlaProAlaSerValSerGlyLysAlaAspGlySerThr 500
Db 1698 GAGCTACGACAGTGCAGAGTGGCCCGCAGCTCTCAGGGCAGAGCGCCGAGGCTTCAC 1757
Qy 501 SerProArgSerGlnGluPro 507
Db 1758 TCGCCCGGAGAGCCGAGAGCC 1778

RESULT 5

US-08-604-989A-10
Sequence 10, Application US/08604989A
Patent No. 5834208

GENERAL INFORMATION:

APPLICANT: Sakano, S.

TITLE OF INVENTION: No. 5834208a1 Tyrosine Kinase

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/604,989A

FILING DATE: February 23, 1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Charles E. Miller

REGISTRATION NUMBER: 24,576

REFERENCE/DOCKET NUMBER: 1920-026

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1521 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: human

STRAIN: UT-7

US-08-604-989A-10

Alignment Scores:

Pred. No.: 1.17e-256
Score: 2662.00
Percent Similarity: 99.808
Best Local Similarity: 99.808
Query Match: 99.668
DB: 2
Gaps: 0

US-09-977-260-2 (1-507) x US-08-604-989A-10 (1-1521)

Qy 1 MetLlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
Db 1 ATGGGGGGGAGGCTCTGTGTCTCTGGGGGCAATTCACGGCTGTGATCTGTGTAG 60
Qy 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
Db 61 GAATTCCTCCGGGAGACCCCGCTCTCGAGACCTCGGACCCCGCTCCGCTCGAGCC 120
Qy 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnHisThr 60
Db 121 AGGATGCCAACGAGGCGCTGGGCCCGCCGAGACCCAGTGTATCACCAATCGAGACACC 180
Qy 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
Db 181 CGCCCCAGCCAGGGGAGCTGGCTTCCTCCGAGGGGCGACGTGTACCATCTGGAGGCC 240
Qy 81 CysGluAsnLysSerTrpTyArgValLysHisThrSerGlyGlnGluLeu 100
Db 241 TCGGAGAAACAAGAGCTGTACCGGTCACACACACAGTGAGAGAGGGGCTGTCTG 300
Qy 101 AlAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 301 GCACTGGGGCGCTGGCGGAGCGGAGGCCCTCTCCGACAGCCCAAGCTCAGCTCATG 360
Qy 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaGlnGlnLeuGlnProProGlu 140
Db 361 CCGTGTGTCACGAGGAAGATCTGGGCGAGAGAGCTGTCCAGAGCTGACGCTCCGAG 420
Qy 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyValLeuCysVal 160
Db 421 GATGGCTGTCTCTGGTGGCGGAGTCCGCGCCAGCCCGGAGCATACGCTGTGCGTG 480
Qy 161 SerPheGlyArgAspValIleHisTyArgValLeuHisArgAspGlyHisLeuThrIle 180
Db 481 AGCTTTGGCCGCGCATCATCACTACCGGTCTGTGAGCGGACGCCACCTCACAAATC 540
Qy 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTySerLysAspLys 200
Db 541 GATGAGCGCGCTGTCTTCTGCAACCTCATGAGCATGTGAGCATTAACAGAGCAAG 600
Qy 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
Db 601 GCGGCTATGTGCACCAAGCTGTGAGAACCAACGAGAAACACGAGCAAGTGGCGGAG 660
Qy 221 GluGluLeuAlaArgAlaGlyTrpPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
Db 661 GAGGAGCTGGCCAGGGCGGGCTGTACTGAACCTGCACGATTGACATTGGGAGACAG 720
Qy 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyGlyTyTrpLeuGlyGlnLysValAla 260
Db 721 ATCGAGAGGGAAGATTGGAGCTGTCTGACAGGAGTGAAGTGGGCAAAAGTGGCC 780
Qy 261 ValLysAsnLleLysCysAspValIleAlaGlnAlaPheLeuAspLutlralAlaMet 280
Db 781 GTGAAGAAATATCAAGTGTATGTGACAGCCAGCGCTTCTGTGACAGAGAGCGGCTCAG 840
Qy 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
Db 841 ACGAAGATGCACACAGAACCTGGTGTCTCTGCGGTATCTTGACACAGGGGCTG 900
Qy 301 TyrlleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgTrpArgGly 320
Db 901 TACATGTGATGAGACAGAGAGCAAGGAGCAAGCTGTGACATTGTGGAGACCGGGGT 960
Qy 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGlyMet 340
Db 961 CGAGCCCTGTGAAACCCCTCAGCTCTGCAAGTTTCTCTGACGCTGCGCGGAGCATG 1020
Qy 341 GluTyrlleGluSerLysLysLeuValHisArgAspLeuAlaArgAsnLleLeuVal 360

Db 1021 GAGTACCTGGAGACCAAGACCTTGTCACCGCAGACCTGGCCGCCGCCACATCTCGTGC 1080
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGlyArgLysGly 380
Db 1081 TCAGAGGACCTGTGGCCAGAGGTGACGACTTGTGGCTGGCCAAAGCCGAGAGGGGG 1140
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
Db 1141 CTAGACTCAACGCGCGCTGCCCTCAAGTGAAGGCGCCGAGGCTCTCAACACGAGGAG 1200
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420
Db 1201 TTCACGACGACAGTGGATGTGTGAGTTTGGGGTGTGCTGCTGGAGCTCTCTCAAT 1260
QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGlyValSerGluAlaValGluLysGly 440
Db 1261 GGAGGGGCTCCGTACCTAAATGTCACTGAAGAGGTGTGAGGCGGTGAGAGGGG 1320
QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerCys 460
Db 1321 TACCGCATGAACCCCGCAGAGGCTGTCCAGGCCCGCCGACGTCATAGAGCAGTGC 1380
QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db 1381 TGGGAGGACAGCCCGCCCGCCGACCTTCCGCAACCTGGCCGAGAACCTGGCCGG 1440
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyHisAspAlaAspGlySerThr 500
Db 1441 GAGCTACGCACTGCAGGTGCCCCCAGCTTCGCTCAGGGGAGAGGCGGAGGCTCCACC 1500
QY 501 SerProArgSerGlnGluPro 507
Db 1501 TCGCCCGAAGCAGAGAGCC 1521

RESULT 6
US-08-876-882-1
; Sequence 1, Application US/08876882
; Patent No. 5981201
; GENERAL INFORMATION:
; APPLICANT: Avraham, Hava
; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT
; TITLE OF INVENTION: OF BREAST CANCER
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds P. C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/876,882
; FILING DATE: 16-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/035,228
; FILING DATE: 08-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Doreen, Hogle M
; REGISTRATION NUMBER: 36,361
; REFERENCE/DOCKET NUMBER: NEDH97-01PA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 781-861-6240
; TELEFAX: 781-861-9540
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1987 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-876-882-1

Alignment Scores:
Pred. No.: 2,02e-249 Length: 1987
Score: 2591.50 Matches: 502
Percent Similarity: 99.21% Conserved: 1
Best Local Similarity: 99.01% Mismatches: 3
Query Match: 97.02% Indels: 1
Gaps: 3

US-09-977-260-2 (1-507) x US-08-876-882-1 (1-1987)

QY 1 MetaLagIyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
Db 263 ATGGCGGGGCGAGGCTCTGTGGTTCCGTGGCGGCATTTCACGGCTGTGATTCTGTGAG 322
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla 40
Db 323 GAACCTCCCGGGGTAGCCCGCCCTTCCCGAGCTTGACACCCCTCCGCTCAGCC 382
QY 41 ArgMetProTrpArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
Db 383 AGGATGCCAAGAGGCGCTGGGCCCCGGGACCCAGTGTATCCCAATTCGAGACACAC 442
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
Db 443 CGCCCAAGCCAGGGAGCTGGCTCCGCAAGGGCGACGTGTCAACCATCTCGAGAGGCC 502
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisIleThrSerGlyGlnGlyLeu 100
Db 503 TCGGAGAACAAAGCTGTACCGCTCAAGACCAACACAGTGAAGAGAGGGCTGTG 562
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 563 GCAGCTGGGGGCTGGGAGAGGGAGGCCCTCTCCGAGAACCCCAACCTCAGCTCATMG 622
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
Db 623 CCGTGGTTCCACAGGGAATCTCGGGCCAGAGGCTGTCCAGACCTCAGCTCCGAG 682
QY 141 AspGlyLeuPheLeuValArgGluSerLarArgHisProGlyAspTrpValLeuLysVal 160
Db 683 GATGGCTGTCTGTGTGGGAGTCCCGGCGCCACCCCGGAGTACGTCTGTGCTG 742
QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
Db 743 AGCTTTGGCCCGCAGAGTATCCACTACCGCTGTGACACCCGAGCGCCACCTCAATC 802
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspValGlnHisTrpSerLysAspLys 200
Db 803 GATGAGGCCGTTGTTCTTGCACCTCATGACATGAGTGGAGCATTCACGAGACAG 862
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
Db 863 GGCGCTATCTCCACCAAGCTGTGTGAGACCAAGCGGAAACACGGGACCAAGTGGCCAG 922
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
Db 923 GAGAGAGTGGCCAGGGCGGGCTGTACTGAACCTGCGACATTCATTCAGTTGGAGACAG 982
QY 241 IleGlyGluGlyLysPheGlyAlaValLeuGlnGlyLysIleLysGlnLysValAla 260
Db 983 ATCGAGAGGAGAGATTGGAGCTGTCTGCAGAGGTGAGTACCTGGGCGCAAAAGTGGCC 1042
QY 261 ValLysAsnIleLysCysAspValTrpAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
Db 1043 GTGAAGAATATCAAGTGTGATGTGACACCCAGGCTTCTGTGAGAGAGAGCGGCTCATG 1102
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300

```
Db 1103 ACGAAGATGCACACGAGAACCTGTGCTCTCTGGGCGGTGATCTGACACAGGCGGTG 1162
QY 301 TTTTleValMetGluHisValSerLysGlyAsnLeuValAsnPhenLeuArgGly 320
Db 1163 TACATTTGATGAGCAGCAGTGCAGGAGGCAACCTGGTGAACCTTTCGGCGACCGGGGT 1222
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPhenSerLeuHisValAlaGlyMet 340
Db 1223 CGAGCCCTCTGTAACACCGCTCAGCTCTCTGAGTTTCTCTGACGCTGGCGAGGCGCATG 1282
QY 341 GluTyrLeuGlnSerLysLysValHisArgAspLeuAlaAlaAsnIleLeuVal 360
Db 1283 GAGTACTGGAGACCAAGAACTGTGTACACCGCAGCTGGCGCCCGCCACATCTGTGTC 1342
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGlyArgGly 380
Db 1343 TCAGAGAGACCTGGTGGCCAGAGTCAGCAGCTTGTGCTGGCGCAAGCGAGCGAGGGG 1402
QY 381 LeuAspSerSerArgLeuProValLysTyrThrAlaProGluAlaLeuLysHisGlyLys 400
Db 1403 CTGAGCTCAAGCGCGCTGCGCTCAAGTGAAGCGCGCCGAGGCTCTCAACACAGCGG-- 1459
QY 401 PheThrSerLysSerAspValTyrSerPheGlyValLeuLeuTyrGluValAlaPheSerTyr 420
Db 1460 TTTCACCAAGCAAGTCGATGTCTGAGTGTGGGGTGTGCTGTGGAGGCTCTTCTCATAT 1519
QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
Db 1520 GAGCGGGCTCTGACCTTAAGTCACTGAAGAGGTGTGAGGCGCGTGGAGAAAGGGG 1579
QY 441 TyrArgMetGluProProGluGlyCysProGlyProAlaHisValLeuMetSerSerCys 460
Db 1580 TACCGCTTGAACCCCGGAGGGGTGTCCAGGCGCCCTGCGACGCTCCATGAGAGCTGTC 1639
QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db 1640 TGGGAGGCAAGCGCGCC--CGCGGCGCACCTTCGCGAACTGGCGAGAAAGCTGGCGCGG 1698
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyLysAlaAspGlySerThr 500
Db 1699 GACTACGACAGTGCAGGTGCGCCAGCGCTCCGTCTCAGGCGAGCAGCGCAGCGTCC-ACC 1757
QY 501 SerProArgSerGlnGluPro 507
Db 1758 TCGCGCCGAGACCGAGAGCC 1778

RESULT 7
US-09-315-928-1
; Sequence 1, Application US/09315928
; Patent No. 6368796
; GENERAL INFORMATION:
; APPLICANT: Avraham, Hava
; APPLICANT: Groopman, Jerome E.
; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
; FILE REFERENCE: BREAST CANCER
; CURRENT APPLICATION NUMBER: US/09/315,928
; CURRENT FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 08/876,882
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: US 60/035,228
; PRIOR FILING DATE: 1997-01-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1987
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (263)...(1846)
US-09-315-928-1
```

```
Alignment Scores:
Pred. No.: 2,02e-249 Length: 1987
Score: 2591.50 Matches: 502
Percent Similarity: 99.21% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 3
Query Match: 97.02% Indels: 3
DB: 4 Gaps: 1

US-09-977-260-2 (1-507) x US-09-315-928-1 (1-1987)
QY 1 MetAlaLysArgGlySerLeuValSerTyrArgAlaPheHisGlyCysAspSerAlaGlu 20
Db 263 ATGGCGGGGAGGAGCTTCTGTTCTCTGGCGGGCATTTGACGCGTGTGATCTGTGAG 322
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
Db 323 GAATTTCCCGGGGAGACCCCGCTTCTCGAGCTGGACCCCGCTCCCGTCTGAGCC 382
QY 41 ArgMetProThrArgArgTrrAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
Db 383 AGGATGCCAAGCAGAGCGCTGGCGCCCGGCGCACCGAGTGTATCACAAATGCGAGCACACC 442
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
Db 443 CGCCCAAGCGAGGAGGAGCTGGCTTCCGAAAGGCGACGTGTGCACATCTGAGAGCC 502
QY 81 CysGluAsnLysSerTyrTyrArgValLysHisIleThrSerGlyGlnGlyLeu 100
Db 503 TGGGAGAACAAAGCTGTGTACCGCGTCAACACACACACACGATGGAGAGGGGCTGTG 562
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 563 GCAGCTGGGCGCTGCGGAGCGGAGCGCCCTCTCCGACAGCCCAAGCTCAGCCTCATG 622
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
Db 623 CCGTGTTCACGAGGAGATCTGGGCGAGAGCTGTCCAGAGCTGCGAGCTCCCGAG 662
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysVal 160
Db 683 GATGGGCTGTCTCGTGGGCGGAGTCCGCGCGCACCCCGCGCATGCTGTGCTG 742
QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
Db 743 AGCTTTGGCGCGCAGCATCTACCTACCGGTCTGTGACCGCGCACCGCCCTCACATC 802
QY 181 AspGluAlaValAlaPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
Db 803 GATGAGCGCGTGTCTTCTCTCAACCTCATGAGATGCTGAGCATTTACACAGAGACAG 862
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
Db 863 GCGCTATCTGCACCAAGCTGTGTGAGACCAAGCGGAAACAGCGGACCAAGTGGCGAG 922
QY 221 GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
Db 923 GAGGAGCTGGCGCAGGCGCGCTGTACTGAACCTCAGCATTTGACATTTGAGAGCAG 982
QY 241 IleGlyGlnGlyGluPheGlyAlaValIleGlnGlyGluTyrLeuGlyGlnLysValAla 260
Db 983 ATCGGAGAGGAGAGATTTGAGAGCTGTCTCTGCAAGGTGAGTACTTGGGCGCAAGGTGGCC 1042
QY 261 ValLysAsnIleLysCysAspValIleThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
Db 1043 GTGAGAGATATCAAGTGTATGTGACAGCCAGCGCTCTCTGACAGAGAGCGGCTCATG 1102
QY 281 ThrLysMetGlnHisLysLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
Db 1103 ACGAAGATGCACACAGAACTGTGCTCTCTGAGGCTGATCTGCAACACAGGCGGTG 1162
QY 301 TTTTleValMetGluHisValSerLysGlyAsnLeuValAsnPhenLeuArgGly 320
```

```

Db      1163 TACATTGTCATGGACAGCTGACGAAGGCAACCTGCTGTAAGTCTTCTGCGGACCCGGGT 1222
QY      321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluMet 340
Db      1223 CGAGCCCTCGTGAACACCGCTCAGCTCCGCGAGTTTCTCTCAGCTGGCCGAGGCGATG 1282
QY      341 GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
Db      1283 GAGTACCTGGAAGCAAGAGCTTGTGCACCCGACCTGGCCGCCGCAACATCTGTCGTC 1342
QY      361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaAlaGluArgLysGly 380
Db      1343 TCAGAGACCTGCTGGCCAGGTCGACGACTTTGGCTGGCCAAAGCGGAGGAGGAGG 1402
QY      381 LeuAspSerSerArgLeuProValLysTyrThrAlaProGluAlaLeuHisGlyLys 400
Db      1403 CTAGACTCAGCCGGCTGCTCCCTCAAGTGAAGCGGCCCGAGGCTCTCAACACGGG--- 1459
QY      401 PheThrSerLysSerAspValTyrSerPheGlyValLeuLeuTyrGluValPheSerTyr 420
Db      1460 TTCACACCAAGTCGATGTCGAGTTTGGGGTGTGCTGCTGGAGGTCCTTCATAT 1519
QY      421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
Db      1520 GGACGGCTCCGTACCTTAATGTCACTGAAGAAGGTCTCGAGGCGGTGAGAGAGG 1579
QY      441 TyrArgMetGluProGluGluLysProGluProValHisValLeuMetSerSerCys 460
Db      1580 TACCCATGGAACCCCGGAGGCTGTCCAGGCCCGCTGACGTCCTCATGACAGCTGC 1639
QY      461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db      1640 TGGGAGGCAAGACCCCGC-CGCGGGCCACCTTCCGCAAACTGGCGAAGGTGCGCCGG 1698
QY      481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaGlySerThr 500
Db      1699 GAGTACCGAGTGCAGGAGGCCACGCTCCGCTCAGGAGGAGCGGACGCTCC-ACC 1757
QY      501 SerProArgSerGlnGluPro 507
Db      1758 TCGCCCGAAGCAGGAGGCC 1778

RESULT 8
US-08-604-989A-9
; Sequence 9, Application US/08604989A
; Patent No. 5834208
; GENERAL INFORMATION:
; APPLICANT: Sakano, S.
; TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/604,989A
; FILING DATE: February 23, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles E. Miller
; REGISTRATION NUMBER: 24,576
; REFERENCE/DOCKET NUMBER: 1920-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741

```

```

;
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1398 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: human
; STRAIN: UT-7
; US-08-604-989A-9

Alignment Scores:
Pred. No.: 6,656-235 Length: 1398
Score: 2444.00 Matches: 466
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 91.50% Indels: 0
DB: 2 Gaps: 0

US-09-977-260-2 (1-507) x US-08-604-989A-9 (1-1398)
QY      42 MetProThrArgArgTyrPAlaProGluTyrGlnCysIleThrLysCysGluHisThrArg 61
Db      1 ATGCCACAGAGCGCTGGGGCCCCGGGCAACCGATGATACCAATGCGAGCACACCGCC 60
QY      62 ProLysProGluGluAlaPheArgLysGlyAspValValThrIleLeuGluAlaCys 81
Db      61 CCCAAGCCAGGGGAGCTGGCTTCGCAAGGGCGACGTGTCCATCTCGTAGGCTGC 120
QY      82 GluAsnLysSerTyrTyrArgValLysHisThrSerGlyGlnGluLeuLeuAla 101
Db      121 GAGAACAGAGCTGTATCCGCTCAAGCACACACAGTGGACAGAGGGCTGCTGGCA 180
QY      102 AlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetPro 121
Db      181 GCTGGGGCGCTGCGGAGCGGAGGAGCCCTCTCCGACAGCCCAAGCTCAGCTCATGCGG 240
QY      122 TrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGluAsp 141
Db      241 TGGTTCACAGAGAAATCTCGGGCCAGAGGCTGTCCAGACCTCAGCTCCAGAGAT 300
QY      142 GluLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGluValSer 161
Db      301 GGGCTGTCTCTGCTGGGAGTCCGCGGCCACCCCGGAGCTACTGCTGTGCTGAC 360
QY      162 PheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIleAsp 181
Db      361 TTTGGCCGCGACGTCATCCACTACCGCGCTGTCACCGCAGCGCCACCTCCATTCGAT 420
QY      182 GluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLysGly 201
Db      421 GAGCGCGTCTCTGCAACCTCATGACATGAGGAGCATTAAGCAAGGACAAAGGCG 480
QY      202 AlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 221
Db      481 GCTATCTCAGCAACACTGCTGTAGAACCAAGCGGAACACGGGACCAATCGCGCAGAGAG 540
QY      222 GluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIle 241
Db      541 GAGCTGGCCAGAGCGCGCTGTACTGAACTGACGAGCTTTGACATTTGGAGACAGATC 600
QY      242 GlyGluGluGluPheGlyAlaValLeuGlnGlyGlyTyrLeuGlyGlnLysValAlaVal 261
Db      601 GGAGAGGAGAGTTTGGAGCTGTCTCGAGGAGTGAGTCTGGGGCAAAAGGTGGCGCTG 660
QY      262 LysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMetThr 281
Db      661 AAGATATCAAGTGTGATGTACAGCCAGGCTTCCTCGAGAGAGAGCGCGCTCATAGCG 720
QY      282 LysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeuTyr 301

```

Db 721 AAGATGACACGAGAACCTGTGCTCTCTCTGCGCTGATCCTGCACAGGGCTGTAC 780
Qy 302 ILVAlmetGluHISValSerLysGlyAsnLeuValAsnPhenLeuArgThrArgGlyArg 321
Db 781 ATGTTCATGAGACGACGAGCAAGGGCAACCTGTGTGAATTTCTGCGAGCCGGGGTTCGA 840
Qy 322 AlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHISValAlaGlnGlyMetGlu 341
Db 841 GCCCTCTGTAACACCGCTCAGCTCTCTCAAGTTTCTGTGACGCTGGCCGAGGGCATGAG 900
Qy 342 TyrLeuGlnSerLysLysLeuValHISArgAspLeuAlaAlaArgAsnIleLeuValSer 361
Db 901 TACCTGGAGAGCAAGACTTGTGCACCGGACCTGGCCGCCCAACATCTCTGCTCA 960
Qy 362 GluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGlnArgGlyLeu 381
Db 961 GAGGACCTGGTGGCCAAAGGTGACGACTTGGCTGGCCAAAGCCGAGCGGAAGGGCTA 1020
Qy 382 AspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHISGlyLysPhe 401
Db 1021 GACTCAAGCCGGCTGCTCAAGTGTGACGAGCGGCCCAAGGCTCTCAACACGGGAAGTTC 1080
Qy 402 ThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyrGly 421
Db 1081 ACCAGCAAGTCGATGCTGAGATTGGGGTGTCTGTGGAGGTCTTCTCATATGGA 1140
Qy 422 ArgAlaProTrpProLysMetSerLeuLysGlyValSerLysAlaGlnGlyTyr 441
Db 1141 CGGGCTCTGTAACCTAAAGTCACTGAAAGAGTGTGCGAGGCGCGGGAAGGGGTAC 1200
Qy 442 ArgMetGluProProGlnGlyCysProGlyProValHISValLeuMetSerSerCysTrp 461
Db 1201 CGCATGTAACCCCGGAGGGCTGTCCAGGCCCGCTGTGACGTCTCATGACAGCTGCTGG 1260
Qy 462 GluAlaGluProAlaArgArgProProPheArgLysLeuAlaGlnLysLeuAlaArgGlu 481
Db 1261 GAGGCAAGCCCGCCCGCGCCGACACCTTCCGCAAACTGGCCGGAAGGTGGCCGGAG 1320
Qy 482 LeuArgSerAlaGlyAlaProAlaSerValSerGlyLysAlaAspGlySerThrSer 501
Db 1321 CTACGCAAGTGCAGATGCTCCAGCTCGTCTCAGGGCAGGACGCCGCTCACCTCG 1380
Qy 502 ProArgSerGlnGluPro 507
Db 1381 CCCGAAAGCCAGGAGCC 1398
RESULT 9
US-09-741-154-1
; Sequence 1, Application US/09741154
; Patent No. 6437110
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, ELLEN M. et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001061
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Human
US-09-741-154-1
Alignment Scores:
Pred. No.: 2,16e-192 Length: 1713
Score: 2020.50 Matches: 410
Percent Similarity: 84.888 Conservative: 11
Best Local Similarity: 82.668 Mismatches: 48
Query Match: 75.658 Indels: 29
DB: 4 Gaps: 6

US-09-977-260-2 (1-507) x US-09-741-154-1 (1-1713)
Qy 23 ProArgValSerPro-----ArgPheLeuArgAlaTrp 33
Db 30 CCCCGAGACAGCGCGGGGTGGCAGAGAACACACAGCTGCAGGGGGCTGGAGAACCA 89
Qy 34 HisProProProValSerAlaArgMetProThrArgArgTrpAlaProGlyThrGlnCys 53
Db 90 CACCCCTCTCC-----TCCCCCAGAGAACTGGAGCGGGGACCTCG--- 134
Qy 54 IleThrLysCysGlnHISThrArgProLysProGlyLysLeuAlaPheArgGlyAsp 73
Db 135 GTTCCGGA-TGC-----TGCAGGACCGCGG 160
Qy 74 ValValThrIleLeuGlnAlaCysGluAsnLysSerTrpTyrArgValLysHISThr 93
Db 161 TTGCTAGGACACATCGAGAGAGAGAGAGAGAAAGTTTCATGGGGTCTGACAGCATCT 220
Qy 94 SerGlyGlnGlyLeuLeuAlaAlaGlyAlaLeuArgGlnArgGlnAla---LeuSer 112
Db 221 AAATG-TCCCTCTCATGACAGGACATTTCCCTCTGAAACGAGAGAGCAGACTTCGGC 279
Qy 113 AlaAspProLysLeu---SerLeuMetProTrpPheHISGlyLysLysSerGlnGln 131
Db 280 GAGGAGCCCGTGTGACAGCAGCAGCTTGTGGTGTCCAGCGAAGATCTGGCCAGAG 339
Qy 132 AlaValGlnLeuGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGlnSerAlaArg 151
Db 340 GCTGTCCAGCAGGTGACGCTCCCGAGAGATGGGCTGTCTGTGGTGGGAGTCCGGCCG 339
Qy 152 HisProGlyAspTyrValLeuCysValSerPheGlyArgAspValIleHISThrArgVal 171
Db 400 CACCCCGGCACTACGCTCTGTGAGCTTGGCGGCGAGCTCATCCACTCCCGGTG 459
Qy 172 LeuHISArgAspGlyHISLeuThrIleAspGluAlaValPhePheCysLeuMetAsp 191
Db 460 CTCGACCGGACAGCGCCACTCACAATGATGAGGCGGTGTTCTTGTCAACCTCATAGGAC 519
Qy 192 MetValGlnHISThrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLys 211
Db 520 ATGTGTGAGCATTAACAGCAAGAGAGGGCGGTATCTGCACCAAGCTGTGTGAGAACCAAG 579
Qy 212 ArgLysHISGlyThrLysSerAlaGlnGlnLeuAlaArgAlaGlyTrpLeuLeuAsn 231
Db 580 CGGAAACAGCGGACCAAGTCTGGCGGAGAGAGACTGCGCGGGCGGTGTACTGAAC 639
Qy 232 LeuGlnHISLeuThrLeuGlnAlaGlnIleGlyGlnGluPheGlyAlaValLeuGln 251
Db 640 CTCGACATTTTGCATTTGGAGGCACAGATCGGAGAGAGAGTTTGAGCTGTCCAG 699
Qy 252 GlyLysLysLeuGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGln 271
Db 700 GGTGAGTACTGGGCAAAAGTGGCGCGTGAAGAAATATCAAGTGTATGTAGACAGCCAG 759
Qy 272 AlaPheLeuAspGlyThrAlaValMetThrLysMetGlnHISGlnAsnLeuValArgLeu 291
Db 760 GCCTCTCTGACAGAGCGCGCTCATGAGAGATCAACACAGAGAACTGTGGCTCTTC 819
Qy 292 LeuGlnValIleLeuHISGlnGlyLeuTyrIleValMetGlnHISValSerLysGlyAsn 311
Db 820 CTGGGCGTATCTGCACCAAGGGGCTGTACATTTGTATGAGACCTGAGCAAGGGCAAC 879
Qy 312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
Db 880 CTGTGTAACCTTCTGTGGAGCCCGGGTCTGACGCTGTGTAACACCGCTCATGCTGTGAG 939
Qy 332 PheSerLeuHISValAlaGlnGlyMetGlyTyrLeuGlnSerLysLysLeuValHISArg 351
Db 940 TTTTCTCTCAGCTGCGCAGAGGAGATGAGTACTGAGAGAGAAAGCTTGTGCACCGC 999
Qy 352 AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371

```

Db      1000 GACCTGGCCGCCCCCAACATCTCTGTCACAGAGACCTGTCGGCCAAAGTCAGCCACTTT 1059
Qy      372 G1yLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391
Db      1060 GGCCTGGCCAAAGCCGAGCGGAGGGGCTAGACATCAAGCCGCGTCGCTCAAGTCGAGC 1119
Qy      392 AlAProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTyrSerPheGly 411
Db      1120 GCGCCGAGGCTCTCAAAACAGGAGATTCACACAGCAAGTCGAGATCTGTGAGATTGTCGG 1179
Qy      412 ValLeuLeuTrpGluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLys 431
Db      1180 GTGCTGCTCTGGAGGCTCTCTCATATGACGGGCTCCGTACCTCAAAATGTCATCTGAAA 1239
Qy      432 GluValSerGluAlaValAlaGluLysGlyTyrArgMetGluProProGluGlyCysProGly 451
Db      1240 GAGGTGTCGAGGCGCTGGAAGAGGGTACCGCATGCAACCCCGGAGGCTGTCCAGGC 1299
Qy      452 ProValHisValLeuMetSerSerCysTrpGluAlaGluProAlaArgArgProPhe 471
Db      1300 CCCGTCACCTCTCTCATATGACAGCTGCTGGAGGACAGGCCCGCCCGCCGACACCTTC 1359
Qy      472 ArgLysLeuAlaGluLysLeuAlaArgGluLeuArgSerAlaGlyAlaProAlaSerVal 491
Db      1360 CGCAAACTGGCCGAGAACTGGCCCGGAGCTACGCAAGTCAGAGTGGCCCGACCTCCGTC 1419
Qy      492 SerGlyHisAspAlaAspGlySerThrSerProArgSerGlnGluPro 507
Db      1420 TCAGGGCAGGACGCGGAGGCTCCACCTGCGCCCGCAAGCCAGAGGCC 1467

RESULT 10
US-09-741-154-3
: Sequence 3, Application US/09741154
: Patent No. 6437110
: GENERAL INFORMATION:
: APPLICANT: BEASLEY, Ellen M. et al.
: TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
: TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
: FILE OF INVENTION: THEROPF
: FILE REFERENCE: C1001061
: CURRENT APPLICATION NUMBER: US/09/741,154
: CURRENT FILING DATE: 2000-12-21
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 16389
: TYPE: DNA
: ORGANISM: Human
: US-09-741-154-3

Alignment Scores:
Pred. No.: 1,038-135 Length: 16389
Score: 1469.50 Matches: 501
Percent Similarity: 21.59% Conservative: 0
Best Local Similarity: 21.59% Mismatches: 6
Query Match: 55.02% Indels: 1819
DB: 4 Gaps: 12

US-09-977-260-2 (1-507) x US-09-741-154-3 (1-16389)
Qy      1 MetaLagIArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
Db      7229 ATGGGCGGCGAGGCTCTCTGTTCTCTGGCGGCGATTTCACGGCTGTGATTCCTGAG 7288
Qy      21 Glu----- 21
Db      7289 GAACCTCCCGGGTAAGATCACTTCCACAGGGGCTTGGGAGACCAGTTCGTGGTCC 7348
Qy      21 ----- 21
Db      7349 ATCCAGATGCTCTGGGAGGCTCTCTGCTGCTCCAGCATGCCCCCACTGACCT 7408
Qy      21 ----- 21

```

```

Db      7409 GGGCTGAAGNGTAGAAGACGACCGTGGGTCCACGCCAAGTCACACCTGACCTCTGT 7468
Qy      22 ---LeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
Db      7469 CCCCTCCCAAGTG-AGCCCCCGCTCTCCGAGCCTGGAGCCACCCCTCCGCTCCAGCC 7527
Qy      41 ArgMetProThr----- 44
Db      7528 AGGATGCCAACGGTAGAGTGTCTAGCCTGCTCTCTCTCCCGGGGTCTCTTCAAC 7587
Qy      44 ----- 44
Db      7588 TCCGTCACACACCTGAGCCCTTCCTCAAGAGGGGTGACCTCTGCCCCCTACCCCGTTC 7647
Qy      44 ----- 44
Db      7648 TGCTGACCTCGGCTGGCCTTTCACATCATTTGGCTCTCTGACCTTCTGTCACAGGCC 7707
Qy      44 ----- 44
Db      7708 GCTCTCTTGTGACTACGATTTCTCTGACACACCCACTCCCTTTCTTTCCCC 7767
Qy      44 ----- 44
Db      7768 GATGTCGCGTCTGTTTTCTCTGCTCCGCTCTTTTCAATCTATACCTTTCTCC 7827
Qy      44 ----- 44
Db      7828 TCCTGCTCTTTCCTCCCTCCGCTCTGCTCCACATCTCGTAGTCTCTCTT 7887
Qy      45 ----- 44
Db      7888 GCCCTCTCCCGGCCCTCTGCAAGAGCGCTGGGCCCGGACCCAGTATACACCA 7947
Qy      56 scYsGluHisThrArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAla 76
Db      7948 ATGGAGACACACCGCCCAAGCAGGGAGACTGCGCTCCGCAAGGGCGAGCGTCCAC 8007
Qy      76 r1LeuGluAlaLysGlu----- 82
Db      8008 CATCTGAGGCTGCGCA-GGTGAGAGTGGCGCGGGTGTGGTGTGGGCTGGGG 8066
Qy      83 -----As 83
Db      8067 GCTCCACGAGACACACCCACACCCCACTAACCCTGCTCTCTCTGCGACAA 8126
Qy      83 nLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluLysLeuAlaAla 103
Db      8127 CAAGAGCTGTACCGCGTCACAGCACACACCATGAGAGAGGGCTGTGGCAGCTGG 8186
Qy      103 yAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetPro----- 121
Db      8187 GCGCTGGGAGGAGGGAGGCGCTCTCCGAGACCCCAAGCTCAGCTGAGT 8246
Qy      121 ----- 121
Db      8247 GGGCAGACAGGGGCTGGGCTAGGGAGACAGTACCCCTCCACAGCCAGCTCT 8306
Qy      122 -----TrpHisGlyLysLysSerGlyGlnGluAlaVal 133
Db      8307 GACCCACCCCTTCGCTGGCGCCAGAGTGCTTCACAGGGAATCTCGGGCCAGAGGCTGT 8366
Qy      133 lGlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGluSerAlaArgHis 153
Db      8367 CCAGCAGCTGACCTCCCGAGATGGGCTCTCTCTGAGGAGAGTCCGCGCCACACC 8426
Qy      153 oGlyAspTrpValLeuCysValSerPheGlyArgAspValLysHisTyrArgValLeu 173
Db      8427 CGGCGACTACGCTCTGCTGCTGAGCTTTGGCGGAGACTATCCACTACCGCGTGTCA 8486
Qy      173 sArgAspGlyHisLeuThrLysAspGluAlaValPhePheCysAsnLeuMetAspMet 193

```

D8487 CCGGAGGCGCCACTCACAATCGATGAGCGCTTCTTCTGCAACCTCATGACATGCT 8546
QY 193 Lglu----- 194
D8547 GGAGGTGCTCCACCCAGAGGCCCCACCCGTCCTGCAATGAGGAGCCAGAGACTGC 8606
QY 194 ----- 194
D8607 GGAGACTCCCTACGTAGAGATGAGAGGGGGGTCACTGACCTGCATCCCTTCCGAGC 8666
QY 194 ----- 194
D8667 AGCTGGGACAGACACATCCCTACCTGCTGTGATCACCAGGAGCCCTGAGCTCTGT 8726
QY 194 ----- 194
D8727 GCTTCCCAATGTGAGATGAGAGGGGTATTCCCAAGGCCCCCAAGGACCCCAAGGCC 8786
QY 194 ----- 194
D8787 CCCATCCCATCTCAGATGGGACAGACAGAGAGACCCCAAGTAGATTCCTTAATGCA 8846
QY 194 ----- 194
D8847 GAAATAGGGGGGCTGATCCCTATGAGACTGCCAATCTACAGGTGGCTGAGTCAAGCTG 8906
QY 194 ----- 194
D8907 TTCTACCTCCAGGAGCTAGAGTACTCTCTGAGGAGACCCCAAGAGACCCCAAGTTGAT 8966
QY 194 ----- 194
D8967 ACCAATCCAGAGACTCCCTCTCTTGTGGCTTGCTCTTCCGCTCTCATGAAACC 9026
QY 194 ----- 194
D9027 AGCCTTCT 9086
QY 195 ----- 195
D9087 AGGATCTGGGGTTCCTCCCTGCTGGGGCTGAGGCTGCTCCACATCTCTGAGCA 9146
QY 195 ----- 195
D9147 TTTCT 9206
QY 215 yThrIysSerAlaGluGluLeuAlaArgAla----- 226
D9207 GACCAAGTGGCGCGAGAGAGAGCTGGCCAGGGG-TAGGGAGCGCCAGAGAGGCAAGCC 9265
QY 226 ----- 226
D9266 CTTTCCCTACCCAGTTAGCCAGTCCGGGAAAGAGGCGCTGGGGCCCGCACCTCTG 9325
QY 226 ----- 226
D9326 AGGCCAGATCACCAAGCTGTCTCATGCCAGCTATGAGATAGAAGACCTGGGCTGCC 9385
QY 226 ----- 226
D9386 TCGGGGGTGGCCACAGACACAGATCAGTCTTATTCATAGTCCCGACCTTGGCAGCT 9445
QY 226 ----- 226
D9446 CATCTGTACATCATCACAATCTGAGCCAGAAACCTGGCCTGAACCTCCACCCACCC 9505
QY 226 ----- 226
D9506 CATCTGTACAGATTTCTCTCTGGCCAGGTTCCCAAGACCTGTGATCTTCTCCCTCC 9565
QY 226 ----- 226
D9566 CCTCCCCACAGCCCAAGCTCAGAGCCCTCCCATCAGACCGGCTCTTCCCAAGAGGGGAGT 9625

QY 226 ----- 226
D9626 CCCCTTCCAGCCAGGCGCTTGTAGTCAGCTCAAGTTATCTTCTATCATGAACCAT 9685
QY 226 ----- 226
D9686 GCCTCCAGGAGAGAGAGGCCAGCAGCTCCACACCTTCCAGAACTGCTGCATGGCTG 9745
QY 226 ----- 226
D9746 AGCCACCTTTCAGGCTCATCTGCAAAAGCCCTTCTCCACACATCCACCCCTCCAGGT 9805
QY 226 ----- 226
D9806 CAATGTGATCATGTTTCTTCTTCATGCTCTGAGCCATTGTATATTCTGTCCCTTACC 9865
QY 226 ----- 226
D9866 TGAATGCTTTTCCAGCCCTTAAGTAGTGAATCTCAATATTAAAGTCAAAAGCAAT 9925
QY 226 ----- 226
D9926 GCCTCCTTGACCCCATGCTGTCTCTCTTGAAGTATCTCTGCTGGGAAAGTCTGTAC 9985
QY 226 ----- 226
D9986 CCATCAGGATGGCATCCAGGTGAGAGACTAGCATCACCTGGCCATGATCTTTAAGAC 10045
QY 226 ----- 226
D10046 CGTGGATTTTGGAGGGGCACTTCTGAGGAGAAAGAGACGCTGCAGAACCGACAG 10105
QY 226 ----- 226
D10106 AGGCGAGTGAAGGGGCAACACAGAGAGTCTTGGAGAAAGGTCAGAGGAGCTAGAGCC 10165
QY 226 ----- 226
D10166 TGAATATTTCTGGAACCTTATAGGCTCCATGTGTGTTTTTCTGCTGTGTGTTT 10225
QY 226 ----- 226
D10226 CTGGGGTGTGAGAGTAAAGAACTGTGAGGAGAGCCAGAGGGTATTGGAGGGAC 10285
QY 226 ----- 226
D10286 AAGAGAGCCAAAGAGAAAGGCGAGCTTATCTCTGAGAACTGAAGTCTTCTGCGAC 10345
QY 226 ----- 226
D10346 TACCCCTTGGGCTGACTGGAGTTCTTGAGGGCAAGTGTGGGCTGAGGGTCTCAGC 10405
QY 226 ----- 226
D10406 CAGGAGGCCCCAGCTTGTACTATTACACACACTATCCCTCTGTAAGCCCTTTTGT 10465
QY 226 ----- 226
D10466 TCAGTCAACCTGAGATAGTTCTGTGTTGCAACAACAATAATGTGTACAGAGTGG 10525
QY 226 ----- 226
D10526 GTGTCTGTATATGTTCTAGGGGCAATGGCCAGGCCAGCCACTCAATGTGTATCCA 10585
QY 226 ----- 226
D10586 CAGCTGCAGACTGTGAGACATCACCTAGTCTCTCACCAAGTAGAAGCCCAATGGGA 10645
QY 227 ----- 230
D10646 GGAATTTAGGGGGCCATTAACCCCTGTGACTGGGTCTTCACTGCACAGCGGGCTGGTACT 10705

Dh 12864 CGCAGCGGAGAGCCCAACATCCCGGGCCATCTCCACCCCAACAGGGCAACCTGT 12923
Qy 313 lAsnPhenLeuArghrArghrArglyAlaValAsnThrAlaGlnLeuGlnPheSe 333
Dh 12924 GAACCTTCTGGGACCCCGGGGTGAGCCCTGTGTAACACCGCTGAGCTCTGCAAGTTTC 12983
Qy 333 rLeu----- 334
Dh 12984 TCTGTAACTGGGACTCTCAGGGTCTGCGGCACCTGGGGTCTGCGGGGCGCAGATTCCAG 13043
Qy 334 ----- 334
Dh 13044 ATCCATCAAGGGGAAACTGAGGACAGGGGCGAGGAGGCTTTGTGAGCTTGTGGCCCAAG 13103
Qy 334 ----- 334
Dh 13104 CACCCCGAGCTTCTGAGCCCTGAAATGGGCAATCTGGACACTGAGCCCCCACTGCCCC 13163
Qy 335 -----HisValAlaGlnGlyMetGluTyrLeuGlnSerLysLysLeuValH 350
Dh 13164 CTACTACCCCGAGCGAGCTGGCGCGAGGCGATGAGTACTGTGAGAGAGAGAAAGCTTGTGC 13223
Qy 350 lAsrAspLeuAlaAlaArghrAsnIlleuValSerGluAspLeuValAlaLysValSera 370
Dh 13224 ACCGCGACTGGCGCGCGCAACATCCTGGTCTCAGAGGACCTGTGGCCCAAGGTCAAGC 13283
Qy 370 sPrPheGlyLeuAlaAlaLysAlaGlnArghrLysGlyLeuAspSerArgLeuProValLysT 390
Dh 13284 ACTTGGCTGTGGCCAAAGCGGAGCGGAGGCGTAACTCAAGCGGCTGCCCTCAAGT 13343
Qy 390 rPrThrAlaProGluAlaLeuLysHsGly----- 399
Dh 13344 GGAGGGGCGCGAGGCTCTCAACACGCGGGGTGAGCCCTGCCCTTCACTACCCCTGGGGCT 13403
Qy 399 ----- 399
Dh 13404 TTGGGGTCCCCCAGCTCTGCTGATGACCTGGGTATGTGCCCTTGGCTCTGTGAACCT 13463
Qy 399 ----- 399
Dh 13464 CCAAGGCAATGGCTATGCTCTCCCAAGAACTTGTGGCCAAATTGTCTCTGTAGCC 13523
Qy 399 ----- 399
Dh 13524 CCCCCTGGGCTCAGTTTCTCCAGCTCTGAAAAAGCGTGGGCTCAGACTGAAGGAA 13583
Qy 399 ----- 399
Dh 13584 GAATTAACCAAGCTTCTCATGCACTTAAGGCTGAGCTAGACTCAACTGGGGGGCGGCTTG 13643
Qy 399 ----- 399
Dh 13644 GGGGAGACTCAATTAGGAGAGAGACAGAGAGTGGGGTCAACCCAGAGAGGCTTCTTG 13703
Qy 399 ----- 399
Dh 13704 AGGAGAGGGGCTGGAGCAGAAACCTGAGGGGCTCCCTCTCAACCCCTCTGGGGCC 13763
Qy 400 -----LysPheThrSerLysSerAspValTrrSerPheGlyValLeuLeuTrrPgluValP 418
Dh 13764 CACAGAGATTACACACAGAGTGTGAGATTGTTGGGAGTGTCTGGAGAGGTCT 13823
Qy 418 hSerTyrGlyArghrAlaProTyrProLysMetSer----- 429
Dh 13824 TCTCATATGGACGGGCTCGTACCTTAATATGT- GAGGGGGGTCCAGAGGAGCACTG 13882
Qy 429 ----- 429
Dh 13883 GGTTCGGGCGAGGTCCAGAGGCTGTGGCCCTGACCCCTGCGCACGCTGTGCGCACT 13942
Qy 430 --LeuLysGluValSerGluAlaValGluLysGlyTyrArghrMetGluProProGluGlyC 449
Dh 13943 CACTGAAGAGGTGTGTGGAGGCGCTGAGAAAGGGGTACCGCATGAAACCCCGGAGGGCT 14002

Qy 449 ysProGlyProValHisValLeuMetSerCysTrrPgluAlaGluProAlaArghrP 469
Dh 14003 GTCCAGGCGCCGTGACAGTCTCTCATGAGACACTGCTGGAGAGAGAGCCCGCGCGGC 14062
Qy 469 rOrProPheArghrLysLeuAlaGlnLysLeuAlaArghrLysSerAlaGlyAlaProA 489
Dh 14063 CACCTTTCGCGAAACTGGCCGAGAGCTGGCGCGGAGACTACGCACTGCAAGTCCCGCAG 14122
Qy 489 lAsrValSerGlyGlnAspAlaAspGlySerThrSerProArghrSerGlnGluPro 507
Dh 14123 CTTCTGCTCAGGCGAGAGCGCGACGCTCTCACTGCGCCCGCAAGCAGAGGCC 14178
RESULT 11
US-08-604-989A-8
Sequence 8, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ. ID NO.: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-8
Alignment Scores:
Pred. No.: 8, 81e-118
Score: 1269.00
Percent Similarity: 100.008
Best Local Similarity: 100.008
Query Match: 47.518
Gaps: 0
US-09-977-260-2 (1-507) x US-08-604-989A-8 (1-738)
Qy 233 GlnHisLeuThrLeuGlyAlaGlnIlleGlyGlnGlyGluPheGlyAlaValLeuGlnGly 252
Dh 1 CAGCATTTGACATTGGGAGACAGATCCGAGAGGAGAGACTTGGAGCTCTCTCGCAGGCT 60
Qy 253 GlnTyrLeuGlyGlnLysValAlaValLysAsnIlleLysCysAspValThrAlaGlnAla 272

Db 61 GAGTACCTGGGCGAAGAGTGGCCGCTGAAGAAATATCAAGTGTGATGTGACAGCCGACGCC 120
Qy 273 PheLeuAspGluThrAlaValMetThrIysMetGlnHisGluAsnLeuValArgLeuLeu 292
Db 121 TTCCTGGAGAGACGGCCCTCATGACGAGATGCAACGACGACCACTGGTGGCTCTCCG 180
Qy 293 GlyValIleLeuHisGlnGlyLeuTyrIleValMetGlnHisValSerIysGlnAsnLeu 312
Db 181 GCGGTGATCTCCACACGAGGGCTGTACATTTGATGAGACAGCTGAGCAAGGACCACTG 240
Qy 313 ValAsnPheLeuArgThrArgIlyArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPhe 332
Db 241 GTGAACCTTCTCGGAGCCGGGGCTGAGCCCTGCGAAGACCGCTCAGCTCTGAGTTT 300
Qy 333 SerLeuHisValAlaGlnIlyMetGluTyrIleGlnIlySerIysLeuValHisArgAsp 352
Db 301 TCTGTGACAGTGGCGGAGGAGCATGGATGCTGAGAGACGAAAGCTTGTGACCGCGAC 360
Qy 353 LeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaIlyValSerAspPheGly 372
Db 361 CTGGCCGCGCCGCAACATCTGCTCAGAGACCTGGTGGCAAGGTGACGACTTTGGC 420
Qy 373 LeuAlaIlyAlaGlnArgIlyLeuAspSerSerArgLeuProValIlyStrphAla 392
Db 421 CTGGCCAAAGCCGAGCGGAGGGCTAGACTCAAGCGGCTCCGCTCAAGTGGACGGCG 480
Qy 393 ProGluAlaLeuIlyHisGlyIlySerPheThrSerIysSerAspValIlyTrpSerPheGly 412
Db 481 CCCGAGGCTCTCAACACGGGAGAGTTCACACGACAGTGGAGTCTGGAGTTTGGGGTG 540
Qy 413 LeuLeuTyrPgluValPheSerTyrGlyArgAlaProTyrProIysMetSerLeuIysGlu 432
Db 541 CTGCTCTGGAGGTCTTCTCATATGACGCGGCTCCGTCAAAATGTCATGAAAGAG 600
Qy 433 ValSerGluAlaValGluIlySerIlyArgMetGlnProProGlnIlyCysProGlyPro 452
Db 601 GTGTGGAGCGCCGTGAGAGGGGTACCGCATGGAACCCCGAGGGCTGTCCAGCGCCC 660
Qy 453 ValHisValLeuMetSerSerCysTyrPgluAlaGluProAlaArgProProPheArg 472
Db 661 GTGACGTCCTCATGAGAGCTGTGGAGGACGAGCCCGCGCGGACACCTTCCCG 720
Qy 473 LysLeuAlaGluIlyLeu 478
Db 721 AAACGTGCCGAGAACTG 738
RESULT 12
US-07-820-011A-1
Sequence 1, Application US/07820011A
PATENT No. 5336615
GENERAL INFORMATION:
APPLICANT: Bell, Leonard
APPLICANT: Madril, Joseph A.
APPLICANT: Warren, Stephen L.
APPLICANT: Luthringer, Daniel J.
TITLE OF INVENTION: Genetically Engineered
TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced
TITLE OF INVENTION: Migration
TITLE OF INVENTION: and Plasmidogen Activator Activity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb storage
COMPUTER: IBM PC XT
OPERATING SYSTEM: PC-DOS/MS-DOS 2.10
SOFTWARE: Displaywrite 3
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/820,011A
FILING DATE: 19920106
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: LB-101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1602 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gallus, gallus
PUBLICATION INFORMATION:
AUTHORS: Takeya, Tatsuo
AUTHORS: Hanafusa, Hidesaburo
TITLE: Structure and Sequence of the
TITLE: Cellular Gene Homologous to the RSV src
TITLE: Gene and the Mechanism for Generating the
JOURNAL: Cell
VOLUME: 32
PAGES: 881-890
DATE: March, 1983
US-07-820-011A-1
Alignment Scores:
Pred. No.: 4,51e-64
Score: 737.50
Percent Similarity: 52.84%
Best Local Similarity: 36.40%
Query Match: 27.61%
DB: 1
Gaps: 15
US-09-977-260-2 (1-507) x US-07-820-011A-1 (1-1602)
Qy 23 ProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla----- 40
Db 137 CCCCGAGCGGCTCT-----TTGGACCGTGGCCAGCCGCAAGCTCTTGGGGGCT 190
Qy 41 -----ArgMetProThrArgArg-----TrrAlaProGly--- 50
Db 191 TCACACTTGTGACACCGCTTACGTGCGCCGACGCGCGGGGCACTGGCTGGC-GGCGTC 249
Qy 51 ThrGlnCysIleThrIlyCysGlnHisThrArgProIlyProGlyGluLeuAlaPheArg 70
Db 250 AACACTTGTGCGCTCTACGACACGATGCCGAGTCCGAGTGAACGAGCTTGTCTTCAAG 309
Qy 71 LysGlyAspValValThrIleLeuGlnIlyAcylGluAsnIlySerTyrTyrValIlys 90
Db 310 AAAGGAGAACCGCTGACAGATTGTCAACACGGAAGT---GACTGTGCTGGCTCAT 366
Qy 91 HisHisThrSerGlnGlnIlyLeuLeuAlaIlyAlaLeuArgGluArgGluAla 110
Db 367 TCCCTCACTACAGGACAGACGGGCTACATCCCGATACATGTCGCCCTCAGACGCC 426
Qy 111 LeuSerAlaAspProIlyLeuSerLeuMetProTrrPheHisGlyIlyIleSerGlyGln 130
Db 427 ATCCAGGCTGAA-----GAGTGTACTTGGGAAATCATCATCTGCTGG 468
Qy 131 GluAlaValGlnGlnLeuGlnProProGluAsp-----GlyLeuPheLeuValArgGlu 148
Db 469 GAGTCCGAGCGGCTGCTCTCAACCCGAAACCCCGCGGGAACCTTCTTGGCGGGAG 528
Qy 149 SerAlaArgHisProGlyAspTyrValLeuGlnValSer-----PheGly 163

```

Db      529 AGCGAGACGACAAAGGTGCTTATGCTTCCGTTTCTGACATTGCAACAGCCCAAGGG 568
Qy      164 ArgAspValIleHISrYrArGValLeuHISrArgAsp---GlyHISLeuThrIleAspGlu 182
Db      589 CTCATGTGTGAAGACCTACAGATCCGACAGCTGACAGCGCGCTTCTACATCAGCTCA 648
Qy      183 AlAValPhePheCysAsnLeuMetAspMetValGluHISrSerLysAspLysGlyAla 202
Db      649 CGCACACAGCTTACAGACCTGCGACAGCTGTGGCCCTACACTCAACAACTGCTGATGCG 708
Qy      203 IleCysThrLysLeu-----ValArgProLysArgLysHISrLysGlyThrLysSerAlaGlu 220
Db      709 TTGTGCTCCAGCTGACCAAGCTGTGCGACCTCCAGGCCAGCCAGCGGA----- 762
Qy      221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHISrLeuThrLeuGlyAlaGln 240
Db      763 -----CTGCGCAAGAGACGGGTGGAAATCCCGGGAGTCGCGCTGCGAGGTGAG 816
Qy      241 IleGlyGluGlyLupheGlyAlaValLeuGlnGlyLupheGlyLys---GlnLysVal 259
Db      817 CTGGGGCAGGGCTGCTTGGAGAGCTGTGGATGGGAGCTGGAAGCGCACACAGAGTG 876
Qy      260 AlAValLysAsnLys---CysAspValThrAlaGlnAlaPheLeuAspGluThrAla 278
Db      877 GCCATTAAGACTGTGAAGCCCGGCAACATGTCCCGGAGGCTCTCGACGAAAGCCCAA 936
Qy      279 ValMetThrLysMetGlnHISrLysAsnLeuValArgLeuLeuGlyValIleLeuHISrGln 298
Db      937 GTGATGAGAAGACTCGCGCATGAGAGCTGTTCAGCTGACAGCAGTGGTGGGAAGAG 996
Qy      299 GlyLeuThrLysValMetGlnHISrLysSerLysGlyAsnLeuValAsnPheLeuArgThr 318
Db      997 CCATCTACACTCTGCTACTGATGATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1056
Qy      319 ArgGlyArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHISrValAlaGlu 338
Db      1057 GAATGGGCGAAGTACCTCGCGCTGCCACAGCTGTGATATGCTGCTCAGATTGCACTCC 1116
Qy      339 GlyMetGluThrLeuGlnLysSerLysLysLeuValHISrArgAspLeuAlaArgAsnIle 358
Db      1117 GGCATGGGCTATGTGAGAGAGATGACACTGCTGACCGAGACCTGGGGGGGCGCAACATC 1176
Qy      359 LeuValSerLysAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAla----- 376
Db      1177 CTGTGGGGGGAACCTGTGCTGCAAGGTGGCTGACTTGTGGCTGGCAGCCCTCATCAGAG 1236
Qy      377 -----GluArgLysGlyLeuAspSerArgLeuProValLysTrpThrAlaProGlu 394
Db      1237 GACACAGAGTACACAGCAGCGCAAGTGCACAGTTCCTCCATCAAGTGGACAGCCCGGAG 1296
Qy      395 AlaLeuLysHISrLysPheThrSerLysSerAspValTrpSerPheGlyValLeuLeu 414
Db      1297 GCAGCCCTATATGCGCGGTTCACACATCAAGTGGATGTGCGTTCCTGGCATCTGCTG 1356
Qy      415 TrpGluValPheSerLysGlyArgAlaProTyProLysMetSerLeuLysGluValSer 434
Db      1357 ACTGACTGACCAAGGCGCGGTGCCATACCGAGGATGCTCAACAGGAGGTGCTG 1416
Qy      435 GluAlaValGluLysGlyTrpArgMetGluProProGlnLysGlyProGlyProValHIS 454
Db      1417 GACACAGTGGAGAGGGGCTACCGCATGCTCCGCGCGCGAGCCCGAGTGGCTGCAAT 1476
Qy      455 ValLeuMetSerSerCysTrpGluAlaGluProAlaArgProPheArgLysLeu 474
Db      1477 GACCTCATGTGCAAGTGTGGCGGAGGAGACCTGAGAGAGGGCGCACTTTTGTAGTACTG 1536
Qy      475 AlaGluLysLeuAlaArgGluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGln 494
Db      1537 CAGGCTTCTCTG-----GAG 1551
Qy      495 AspAlaAspLysSerThrSerProArgSerGln 505

```

```

Db      1552 GACTACTTACCTGACAGAGCCCGACGTACCAG 1584
RESULT 13
PCT-US93-00445-1
Sequence 1, Application PC/TUS9300445
GENERAL INFORMATION:
APPLICANT: Bell, Leonard
APPLICANT: Madl, Joseph A.
APPLICANT: Warren, Stephen L.
APPLICANT: Lubringer, Daniel J.
TITLE OF INVENTION: Genetically Engineered
TITLE OF INVENTION: Endothelial Cells
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 760 KB storage
COMPUTER: DELL 486/50
OPERATING SYSTEM: DOS 5.0
SOFTWARE: Displaywrite 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00445
FILING DATE: 19930105
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/820,011
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-101PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1602 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gallus, gallus
PUBLICATION INFORMATION:
AUTHORS: Takeya, Tatsuo
AUTHORS: Hanafusa, Hidesaburo
TITLE: Cellular Gene Homologous to the RSV src
TITLE: Gene and the Mechanism for Generating the
TITLE: Transforming Virus
JOURNAL: Cell
VOLUME: 32
PAGES: 881-890
DATE: March, 1983
PCT-US93-00445-1
Alignment Scores:
Pred. No.: 4,51e-64 length: 1602
Score: 737.50 Matches: 186
Percent Similarity: 52.848 Conservative: 84
Best Local Similarity: 36.408 Mismatches: 185
Query Match: 27,618 Indels: 57
DB: 5 Gaps: 15
US-09-977-260-2 (1-507) x PCT-US93-00445-1 (1-1602)

```



```

Db      1360 CTGCTGACTGACCTACACCAAGGAGCGGTGCCCTACCTACCGGATGTGTAACCGCGAG 1419
QY      433 ValSerGluAlaValGluLysGlyTyrArgMetGluProGluGlyCysProGlyPro 452
Db      1420 GGTGGTGGACCAAGGAGGCGGTACCGGATGCCCTGCCCGGAGGTGTCCGAGTCC 1479
QY      453 ValHisValLeuMetSerSerCysTrpArgLysGluProAlaArgProGluProGluArg 472
Db      1480 CTGCAAGACCTCATGTGCGACAGTGTGGCGGAGAGACCTGAGAGCGGCCCTTCGAG 1539
QY      473 LysLeuAlaGluLysLeuAlaArgLysLeuArgSerAlaGlyAlaProAlaSerValSer 492
Db      1540 TACCTGCAGGCTCTCTG----- 1557
QY      493 GlyGlnAspAlaAspGlySerThrSerProArgSerGln 505
Db      1558 ---GAGGACTACTTACGTCACGACGAGCCCGAGTACGAG 1593

RESULT 15
; Sequence 3, Application PC/TUS9300445
; GENERAL INFORMATION:
; APPLICANT: Bell, Leonard
; APPLICANT: Madril, Joseph A.
; APPLICANT: Warren, Stephen L.
; APPLICANT: Lutheringer, Daniel J.
; TITLE OF INVENTION: Genetically Engineered
; TITLE OF INVENTION: Endothelial Cells
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 760 Kb storage
; COMPUTER: DELL 486/50
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: Displaywrite 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00445
; FILING DATE: 19930105
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/820, 011
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: ALX-101PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255 1400
; TELEFAX: (203) 254 1101
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1611
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: Chromosome 20
; PUBLICATION INFORMATION:
; AUTHORS: Anderson, Stephen K.
; AUTHORS: Gibbs, Carol P.
; AUTHORS: Tanaka, Akio

```

```

; AUTHORS: Kung, Hsing-Jien
; AUTHORS: Fujita, Donald J.
; TITLE: Human Cellular src Gene:
; TITLE: Nucleotide Sequence and Derived Amino
; TITLE: Acid Sequence of the Region Coding for
; TITLE: the Carboxy-Terminal Two-Thirds of
; TITLE: pp60c-src
; JOURNAL: Molecular and Cellular Biology
; VOLUME: 5
; ISSUE: 5
; PAGES: 1122-1129
; DATE: May, 1985
; PUBLICATION INFORMATION:
; AUTHORS: Tanaka, Akio
; AUTHORS: Gibbs, Carol P.
; AUTHORS: Arthur, Richard R.
; AUTHORS: Anderson, Stephen K.
; AUTHORS: Kung, Hsing-Jien
; TITLE: DNA Sequence Encoding the
; TITLE: Amino-Terminal Region of the Human c-src
; TITLE: Protein: Implications of Sequence
; TITLE: Divergence among src-Type Kinase
; JOURNAL: Molecular and Cellular Biology
; VOLUME: 7
; ISSUE: 5
; PAGES: 1978-1983
; DATE: May, 1987
; PCT-US93-00445-3

Alignment Scores:
Pred. No.: 5 11e-63 Length: 1611
Score: 727.00 Matches: 188
Percent Similarity: 49.01% Conservative: 83
Best Local Similarity: 34.00% Mismatches: 185
Query Match: 27.22% Indels: 98
DB: 5 Gaps: 15

US-09-977-260-2 (1-507) x PCT-US93-00445-3 (1-1611)
QY      24 ArgValSerProArgPhe-----LeuArgAlaTrpHisProPro----- 36
Db      14 AGACCAAGCCCAAGAGGAGCCAGCGGGCGCCGACCTGAGACCCCGCAGAACCTGC 73
QY      37 -----Pro 37
Db      74 ACGGCGTGGGGGGGGGCTTCCCGCTCGACAGCCCGCAGCAGCTCGGCGG 133
QY      38 ValSerAlaArgMetProThrArgArgTrpAlaPro----- 49
Db      134 ACGGCCACCGGGGCCCAAGCGGCGC-TTGCCTCCCGCGCGCGCCGACCCAGACCTGTC 192
QY      50 -----GlyThrGlnCysIleThrLysCysGluHisThrArgProGly 65
Db      193 GGAAGCTTCAACTCTCGACACCGTCACCTCCCGCAGAGGGGGCGCGCGCGGT 252
QY      66 -----GluLeuAla 68
Db      253 GGAAGTACCACTTGTGGCCCTCTATGACTATGATGAGAGCGAGACAGACTGTCC 312
QY      69 PheArgLysGlyAspValValThrIleLeuGluAlaCysGluAsnLysSerTrpTyrArg 88
Db      313 TTCAGAAAGGCGGCGGTCCAGATGTGTACACACACAGAGGGA---GACTGTGCTG 369
QY      89 ValLysHisHisThrSerGlyGlnGlyLeuLeuAlaAlaGlyAlaLeuArgGluArg 108
Db      370 GCCCACTCGCTCAGACAGACAGACAGCTACATCCCGCACTACGTGCGCCCTCC 429
QY      109 GluAlaLeuSerAlaAspProLysLeuSerLeuMetProThrPheHisGlyLysSer 128
Db      430 GACTCCATCCAGGCTGAG-----GAGTGTATTTTGGCAAGATACAC 471

```

QY 129 GlyGlnGluAlaValGlnGlnLeuGlnProProGluAsp-----GlyLeuPheLeuVal 146
 :::|||||: ::: ||| |||||: ||| |||||: |||
Db 472 AGAGGGAGTCAAGACCGGTACTGCTCATGACAGAGAACCCGAGAGGAGCCCTCTCTGCTG 531
QY 147 ArgGlnSerAlaArgHisProGlyAspTyrValLeuGlyValSer----- 161
 ||||| ||| ||| ||| ||| ||| ||| |||
Db 532 CGAGAAAGTGAACACGAAAGGTGGCTACTGCGCTCTGAGTGTGACTTTCGACACGCC 591
QY 162 PheGlyArgAspValIleHisTyrArgValLeuHisArgAsp---GlyHisLeuThrIle 180
 ||| ::: ||| |||||: ::: ||| ||| ||| |||
Db 592 AAGGCGCTCAACGACGACTACAGATCCGACAGCTCGACGCGGCGGCTTCTACATC 651
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys 200
 ||| ::: ||| ||| ::: ||| ::: ||| ::: ||| |||
Db 652 ACCTCGCCGACCCAGTTCACAGCCCTGACAGCAGCTGGTGGCTTACTCTCAACACGCC 711
QY 201 GlyAlaIleCysThrLysLeu-----ValArgProLysArgLysHisGlyThrLysSer 218
 ::: ||| ::: ||| ||| ||| ||| ||| ||| |||
Db 712 GATGGCGCTGTCCACCGGCTCACACCGGTGGCCCAAGCTCCAAAGCCGACACTCAGGCG 771
QY 219 AlaGlnGlnGluLeuAlaArgAlaGlyTyrLeuLeuAsnLeuGlnHisLeuThrLeuGly 238
 |||||: ||| ::: ||| ::: ||| ||| |||
Db 772 -----CTGGCCAAAGATGCGCTGGAGATCCCTCGGAGATCGCTCGGCTGGAG 819
QY 239 AlaGlnIleGlnGlnGlnGlnPheGlyAlaValLeuGlnGlnGlyTyrLeuGly---Gln 257
 ::: ||| ::: ||| ||| ||| ||| ||| ||| |||
Db 820 GTCAAGCTGGGCGCGGCTGCTTGGGAGGTGGATGGGAGCTGGAAAGGTACCACC 879
QY 258 LysValAlaValLysAsnIleLys---CysAspValThrAlaGlnAlaPheLeuAspGlu 276
 ::: ||| ::: ||| ::: ||| ::: ||| ::: ||| ||| |||
Db 880 AGGTGGCCATCAAAACCCCTGAAGCCCTGGCAGCATGTCTCCAGAGGCTTCTCCAGAG 939
QY 277 ThrAlaValMetThrLysMetGlnHisGlnAsnLeuValArgLeuLeuGlyValIleLeu 296
 ||||| ||||| ||||| ||||| ||||| ||||| ||| |||
Db 940 GCCAGGTCAATGAAGAGCTGAGGATGAGCAAGCTGTCAGTGTATGCTGTGTTCA 999
QY 297 HisGlnLeuTyrIleValIleMetGlnHisValSerLysGlnAsnLeuValAsnPheLeu 316
 ::: ::: ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 GAGAGGCCCATTTTCATCTGTCACGAGATGACAGAGGAGAGTTGCTGACTTCTC 1059
QY 317 ArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisVal 336
 ::: ::: ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 AAGGGGAGACAGGACAGTACGCGGCTGCTGAGTGGTGGATGCGCTCGTCAATC 1119
QY 337 AlaGlnGlnMetGlnTyrLeuGlnSerLysLysLeuValHisArgAspLeuAlaArg 356
 ||| ||||| ||||| ||||| ||||| ||||| ||| |||
Db 1120 GCGTCAGGCGATGGGTAGCGAGCGGATGACACTACCTCCACCGGACCTTCGTGCAGCC 1179
QY 357 AsnIleLeuValSerGluAspLeuValAlaValSerAspPheGlyLeuAlaLysAla 376
 ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1180 AACATCTGTGTGGAGAGAACCTGTGTGTCAAAGTGGCGACTTGGGCTGGCTCGGCTC 1239
QY 377 -----GluArgLysGlyLeuAspSerSerArgLeuProValLysTyrThrAla 392
 ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1240 ATTGAAGACAATGATGACAGCGCGGCGGCAAGTGGCCAAATTCGCCATCAAGTGGACGCT 1299
QY 393 ProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTyrSerPheGlyVal 412
 ||||| ::: ||| ||| ||| ||| ||| ||| ||| |||
Db 1300 CCAAGAGCTGCTCTATGGCGGCTTACCATCAAGTCGACGTCGTGCTTCGGGATC 1359
QY 413 LeuLeuTyrPgluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLysGlu 432
 ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1360 CTGCTGACTGAGTCAACACAAAGGAGCGGCTGCTTACCTGCGATGGTGAACCGGAG 1419
QY 433 ValSerGluAlaValGlnLysGlyTyrArgMetGluProGlnGlnLysProGlyPro 452
 ||| ::: ||| ||| ||| ||| ||| ||| ||| |||
Db 1420 GTGCTGGACCGAGCGGCTGACCGGATGCCCTGCCCGGAGTGTCCGAGTCC 1479
QY 453 ValHisValLeuMetSerSerCysTyrPgluAlaGluProAlaArgArgProProPheArg 472
 ::: ||| ||||| ||||| ||||| ||||| ||||| ||| |||
Db 1480 CTCACAGACTCTATGTGCGCAGTGTGCGGAAAGAGCTGAGAGAGCGGCCACCTTCGAG 1539
QY 473 LysLeuAlaGlnLysLeuAlaArgGluLeuArgSerAlaGlyAlaProAlaSerValSer 492

Db 1540 TACCTGAGGCGCTTCTCTG----- 1557

QY 493 GlyGlnAspAlaAspGlySerThrSerProArgSerGln 505
Db 1558 ---GAGGACTACTTACAGTCTCCACGAGCCCGACGATACCAAG 1593

Search completed: August 2, 2003, 10:07:13
Job time : 316 secs

1:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1980.DAT *
2:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1981.DAT *
3:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1982.DAT *
4:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1983.DAT *
5:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1984.DAT *
6:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1985.DAT *
7:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1986.DAT *
8:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1987.DAT *
9:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1988.DAT *
10:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1989.DAT *
11:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1990.DAT *
12:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1991.DAT *
13:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1992.DAT *
14:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1993.DAT *
15:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1994.DAT *
16:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1995.DAT *
17:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1996.DAT *
18:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1997.DAT *
19:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1998.DAT *
20:	/SDSI/gcgdata/genseq/genseqn-emb1/NA1999.DAT *
21:	/SDSI/gcgdata/genseq/genseqn-emb1/NA2000.DAT *
22:	/SDSI/gcgdata/genseq/genseqn-emb1/NA2001A.DAT *
23:	/SDSI/gcgdata/genseq/genseqn-emb1/NA2001B.DAT *
24:	/SDSI/gcgdata/genseq/genseqn-emb1/NA2002.DAT *
25:	/SDSI/gcgdata/genseq/genseqn-emb1/NA2003.DAT *

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2671	100.0	1997	24	AAI50631	Human matk gene-Fe
2	2668	99.9	1942	16	AAO84888	DNA encoding cytop
3	2668	99.9	2000	16	AAT00616	Megakaryocyte kina
4	2591.5	97.0	1987	19	AAV44497	Human matk cDNA.
5	2020.5	75.6	1713	24	ABR888791	CDNA encoding huma
6	1540.5	57.7	8212	24	ABR99884	Mouse encoding huma
7	1469.5	55.0	16389	24	ABR88792	Gene encoding huma
8	1247.5	46.7	2187	21	AAZ29701	Wild-type human c-
9	1247.5	46.7	2187	22	AAH23358	Nucleotide sequenc
10	1247.5	46.7	2187	24	ABR84602	Human cDNA differ
11	1247.5	46.7	2187	24	ABR65918	Breast cancer relat
12	1247.5	46.7	2420	21	AAZ44489	PKA substrate, Cak
13	1247.5	46.7	2420	25	ACC50120	PKA substrate, Cak
14	1247.5	46.7	2466	22	AAI93504	Breast cancer asso
15	1196.5	44.8	2442	24	AAST73326	Human polynucleoti
16	994.5	37.2	2591	23	AAST73299	DNA encoding human
17	832.5	31.2	14283	23	ABR11798	Drosophila melanog
18	749.5	28.1	2282	23	ABSS6203	Drosophila melanog
19	744	27.9	4535	23	ABSL07083	CDNA encoding tumo
20	738.5	27.6	2032	21	AAZ45491	Drosophila melanog
21	737.5	27.6	1602	14	AAO46687	PKA substrate, Src
22	735.5	27.5	1759	21	AAZ29700	Chicken pp60 c-src
23	735.5	27.5	1759	22	AAH28357	Wild-type chicken
24	727	27.2	1611	14	AAO46688	Nucleotide sequenc
25	727	27.2	1611	24	ABST73325	Human pp60 c-src g
26	727	27.2	1611	25	ABZ59382	DNA encoding human
27	727	27.2	2129	25	ACA56854	Human src-c encodi
28	725.5	27.2	2015	24	ABK83939	Human signaling p
29	725.5	27.2	2015	24	ABL66673	Human cDNA differ
30	725	27.1	1936	24	ABK83940	Lung cancer relate
31	718.5	26.9	1636	25	ABZ55389	Human cDNA differ
32	718	26.9	4466	24	ABN55752	Mouse src-c encodi
33	715.5	26.8	1911	24	ABK63704	Novel human coding
34	710	26.6	4517	20	AAH90200	Rat sequence diffe
35	710	26.6	4517	22	AAH28359	Human yes1 encodin
36	710	26.6	4517	23	AAST4489	Nucleotide sequenc
37	710	26.6	4517	24	ABV94313	DNA encoding novel
38	709.5	26.6	2288	24	ABK83935	Breast carcinoma r
39	704	26.4	3527	23	AAZ56598	Human cDNA differ
40	704	26.4	7487	23	AAAS92457	DNA encoding novel
41	702.5	26.3	4414	24	ABZ35401	DNA encoding novel
42	701.5	26.3	1614	25	ABX74390	Human gene express
43	701.5	26.3	5520	18	ABN61865	Human cDNA sequenc
44	701.5	26.3	5527	24	ABN97185	c-abl gene. Homo
45	701	26.2	3258	24	ABQ74506	Gene #3713 used to
						Mouse Fyn nucleoti
ALIGNMENTS						
RESULT 1						
ID	AAI50631	standard; DNA; 1997 BP.				
XX	AAI50631:					
XX	02-JAN-2003	(first entry)				
XX	DE	Human matk gene-related coding sequence.				
XX	KW	Human; gene; ds: allergic disease; matk gene; atopic dermatitis.				
XX	OS	Homo sapiens.				
XX	FT	Key				
XX	CDS	Location/Qualifiers				
XX	FT	266..1789				
XX	FT	/*tag- a				


```
Db          1766 TCGCCCAAGCCGAGACGCC 1786

RESULT 2
AA084888
ID      AAQ84888 standard; cDNA to mRNA; 1942 BP.
XX
XX      AAQ84888;
AC
XX
DT       25-MAR-2003   (updated)
DT       26-OCT-1995   (first entry)
DE
XX      DNA encoding cytoplasmic tyrosine kinase.
XX
KW      cytoplasmic; tyrosine kinase; blood; cell differentiation;
KM      screening; anticancer agent; ds.
OS      Homo sapiens.
XX
FH      Key
FH      CDS
FT
FT      Location/Qualifiers
FT           208..1731
FT            /*tag= a
FT        /product= cytoplasmic_c_tyrosine_kinase
FT        /note= "see AAR71133"
FT           349..540
FT            /*tag= b
FT        /note= "encodes SH3 domain (see AAR71129)"
FT           571..795
FT            /*tag= c
FT        /note= "encodes SH2 domain (see AAR71130)"
FT           904..1641
FT            /*tag= d
FT        /note= "encodes tyrosine kinase domain (see
FT              AAR71131)"
FT           331..1728
FT            /*tag= e
FT        /note= "encodes N-terminally truncated form
FT              of the enzyme (see AAR71132)"
FN
FN      misc_feature
FE
FE      misc-feature
FE
FE      misc-feature
FE
FE      misc-feature
FE
FE      misc-feature
PE
PD     02-MAR-1995.
PN      WO9506113-A1.
PP
PR     25-AUG-1994;    94WO-JP01411.
PR     25-AUG-1993;    93JP-0210403.
PR     29-MAR-1994;    94JP-0058553.
PA      (ASAH ) ASAMI KASEI KOGYO KK.
PX
PY      Sakano S;
PI
DR      WIJ; 1995-106842/14.
DR      P-FSDB; AAR71129, AAR71130, AAR71131, AAR71132, AAR71133.
PT
PT      Cytoplasmic tyrosine kinase and antibody recognising it - for
PT      screening chemical substances for tyrosine kinase inhibitory or
PT      activating activity for use as cancer therapy
PS
PS      Claim 7; Page 49-50; 58pp; English.
XX
XX      This DNA encodes a cytoplasmic tyrosine kinase which has enhanced
XX      expression in connection with blood cell differentiation. It was
XX      isolated from the human um-7 blood cell line. The DNA sequences and
XX      antibodies raised against the enzyme, are useful for screening agents
XX      for inhibiting or activating activity on the tyrosine kinase, for
XX      use as anticancer agents.
XX      (Updated on 25-MAR-2003 to correct PN field.)
SQ
Sequence 1942 BP; 365 A; 615 C; 651 G; 311 T; 0 other.
```

Best Local Similarity:	100.00%	Conservative:	1
Best Percent Similarity:	99.80%	Mismatches:	0
Query Match:	99.89%	Indels:	0
DB:	16	Gaps:	0

US-09-977-260-2 (1-507) x AAQ84888 (1-1942)

QY	1	MetaIaGlyArgGlySerLeuValSerThrPheAlaPheHisGlyCysAspSerAlaGlu	20
Db	208	ATGGGGGGGCGAGGCTCTCTGGTTTCCTGGCGGGCATTTCCAGCGCTGTATCTCTGTAG	267
QY	21	GluIleuProArgValSerProArgPheLeuArgAlaThrPheIleProProValSerAla	40
Db	268	GAACATCCCGGGGTAGCCCGCCGCTCTCTCCGAGCCTGGCAKCCCTCCCTGTCAGCC	327
QY	41	ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr	60
Db	328	AGGATGCCACGAGGCGCTGGGCCCCGGGGACCCAGCTGATACCAAAATGCCAGCACCC	387
QY	61	ArgProLysProGluGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla	80
Db	388	CGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGCGAGCTGTACACATCTGGAGGCC	447
QY	81	CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluLeuLeu	100
Db	448	TGCCACACAAAGAGCTGGTACCCTCAAGCCACACACAGCAGGAGAGGGCTGCTG	507
QY	101	AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet	120
Db	508	GCAGCTGGGGGGCTGGGAGGGGAGGCCCTCTCCGAGACCCCAAGCTCAGCTCAG	567
QY	121	ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu	140
Db	568	CCGTGGTTCACAGGGAGATGTCCTGGGGCCAGGAGCGCTGTACAGCAGCTGCAGCTCCGAG	627
QY	141	AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal	160
Db	628	GATGGGCGTGTCTGTGTCGGGAGAGCCCGGGCCACCCCGGAGCTAGCTCTGTGGCTG	687
QY	161	SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle	180
Db	688	AGCTTTGGCCCGCAGCTCATCACCACCGCTGCTGCACCGGAGGCCCACTTCACATC	747
QY	181	AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTrpSerLysAspLys	200
Db	748	GATGAGGCGCGTGTCTCTCTGCAACCTCATGGAACAATGGAGCATTAAGACGAAGACAAG	807
QY	201	GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu	220
Db	808	GGCGCTATCTCACCAACACTGGTGTGAGCCAAAGCGGAAACAGGGACCAAGTGGCCGAG	867
QY	221	GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln	240
Db	868	GAGGAGCTGGCCAGGCGGCTGGTTACTGAACCTGCAGCAATTTGACATTTGGAGACACAG	927
QY	241	IleGlyGlnGlyGluPheGlyAlaValLeuGlnGlnGlyLeuGlyGlnLysValAla	260
Db	928	ATCGGAGGGGAGAGTTTGGAGCTGCTCTGCAAGGGTGTACTGCTGGGCAAAAAGGTGGCC	987
QY	261	ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet	280
Db	988	GTGAAGAATATCAAGTGTGATGTGCACACCAGGCTTCCTGTGAGAGACAGAGGCGCTCATG	104
QY	281	ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu	300
Db	1048	ACGAAGATGCAACACGAGAAACCTGTGTCTCTGGGCGTGAATCCACACAGGCGCTG	1107
QY	301	TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly	320
Db	1108	TACATTGTGCATGGAGCACGTGACGAAGGGCAACTGTGTGAACTTTCTCGGAGCCGGGCT	1166
QY	321	ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet	340

```
Db 1168 CGAGCCCTGTGACACCCCTAGCTCTGCACTTTCTCTGACAGCTGGCCGAGGCGATG 1227
QY 341 GUTYrLeuGlsSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
    |||||
Db 1228 GAGTACTGTGAGAGCAAGACTTGTGCAACCGGAGCTGGCCGCCGCAACATCTCGTGC 1287
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
    |||||
Db 1288 TCAGAGAGACTGTGGCCCAAGGTCAACGACTTTGGCCTGGCCAAACCGAGCGAAGGGG 1347
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
    |||||
Db 1348 CTAGACTCAAGCGGCTGCCCTCCCTCAAGTGGACGCGCCCGAGAGCTCTCAACACGAGGAG 1407
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
    |||||
Db 1408 TTCACGACCAAGTCTGATGTCTGAGACTTTGGGGTCTGCTCTGGAGAGTCTTCATAT 1467
QY 421 GlyArgAlaProTyrTrpLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
    |||||
Db 1468 GGACGGGCTCCGTACCCCTAAATGTCAGAAAGAGGTGCGAAGGCGGTGGAAGGGG 1527
QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
    |||||
Db 1528 TACCGATGGAAACCCCGCCAGGGCTGTCCAGGCGCCCGTGCAGTCTCTCATGAGCACTGC 1587
QY 461 TrpGluAlaGluProAlaArgTrpProPheArgLysLeuAlaGluLysLeuAlaArg 480
    |||||
Db 1588 TGGGAGGCGAGAGCCCGCCGCGCCACCTTCGCCAAACTGGCGAAGAGTGGCGCGG 1647
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
    |||||
Db 1648 GAGCTACGACAGTGCAGAGTCCCGCAGCTCTGCTCAGGAGAGACCCGACGAGCTCCACC 1707
QY 501 SerProArgSerGlnGluPro 507
    |||||
Db 1708 TCGCCCGCAAGCCAGAGGCC 1728

RESULT 3
AAT00616 standard; cDNA; 2000 BP.
AC AAT00616;
DT 26-MAR-1996 (first entry)
DE Megakaryocyte kinase MKK1 cDNA.
XX
KW Megakaryocyte kinase-1; MKK1; cytoplasmic tyrosine kinase;
KW cellular signal transduction; leukaemia; myelosis; myelofibrosis;
XX gene therapy; ss.
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 258..1781
    /*tag= a
XX
XX WO9529185-A1.
XX
XX 02-NOV-1995.
XX
XX 24-APR-1995; 95WO-US05008.
XX
XX 21-APR-1995; 95US-0426509.
XX
XX 22-APR-1994; 94US-0232545.
XX
XX (PLNC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (SUGEN-) SUGEN INC.
XX
XX Gishizky M, Sures I, Ullrich A;
XX
XX MPI; 1995-382959/49.
```

```
DR P-PSDB; AAR84181.
XX
XX New poly:nucleotide(s) encoding megakaryocyte tyrosine kinase(s) -
PT used to develop prods. for the treatment and diagnosis of kinase
PT related signal transduction abnormalities.
XX
XX Claim 2: Fig 1A-C; 82pp; English.
XX
XX
XX Overlapping cDNA clones were combined to obtain a sequence (AAT00616)
CC coding for human megakaryocyte kinase MKK1 (AAR84181). The clones
CC were isolated from a fetal brain library using degenerate primers
CC (AAT00614-15) based on conserved regions within the kinase domain of
CC receptor tyrosine kinases. MKK polynucleotides can be used in the
CC prodn. in host cells of recombinant MKK, and in the gene therapy of
CC diseases such as acute megakaryocytic leukaemia, myelofibrosis and
CC acute megakaryocytic myelosis.
XX
SQ Sequence 2000 BP; 375 A; 631 C; 674 G; 320 T; 0 other;

Alignment Scores:
Pred. No.: 2,1e-198 Length: 2000
Score: 2668.00 Matches: 506
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 99.89% Indels: 0
DB: 16 Gaps: 0

US-09-977-260-2 (1-507) x AAT00616 (1-2000)
QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
    |||||
Db 258 ATGGCGGGGAGAGGCTCTGTGGTTCCTGGCGGCAATTCACGGCTGTGATTCGTGTGAG 317
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
    |||||
Db 318 GAACCTCCCGGGGTAGACCCCGCTTCGCCAGCTGGACACCCCTCCGCTTCAGGC 377
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnIsth 60
    |||||
Db 378 AGGATGCCAAGAGGGGCTGGGCGCCGCGGACACCACTGATACCAAAATGCAGACACAC 437
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
    |||||
Db 438 CGCCCGCAAGCGAGGAGCTGCTCCGCAAGGGGAGCTGTCACCATCTCGAGAGCC 497
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisIsthSerGlyGlnGluLeuLeu 100
    |||||
Db 498 TCGGAGAACAAAGCTGTACCGCGTCACAGCACACACAGTGGAGAGAGGCGCTGCG 557
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
    |||||
Db 558 GCAGCTGGGGGCGCTCGGGAGCGGAGGCGCTCTCCGACCCCAAGCTCAGCTCATG 617
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGluLeuLeuProGlu 140
    |||||
Db 618 CCGTGGTTCACGGGAGAAATCTCGGGCCAGAGGCTGTCCACACCTTGAGGCTCCCGAG 677
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
    |||||
Db 678 GATGGGCTGTCTCTGTGGGAGATCCGGGCGCCACCCCGGAGACTACTCTGTGGCG 737
QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisAspArgPlyHisLeuThrIle 180
    |||||
Db 738 AGCTTTGGCGCGAGCTCATCCATACCGCGCTGTGCACCCGAGAGGCGCCACTCAATTC 797
QY 181 AspGluAlaValPhePheCysAsnLeuMetIleGlnIsthSerLysSerLys 200
    |||||
Db 798 GATGAGGCGGCTGTCTGTGCAACTCATGACATGATGTGAGACATTTACAGCAAGGACAAG 857
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
    |||||
Db 858 GGGCGTATCTGACCAAGCTGTGTGAGACCAAGGCGAAACAGGAGCAACATGCGCGAG 917
```

```

QY 221 GIUGLLeuAlaArgAlaGlyTyrPLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
DB 918 GAGGAGCTGGCCAGGGCGGCTGTACTGAACTTCACACATTGGAGTGGAGACAG 977
QY 241 IIEGlyGluGlyPheGlyAlaValLeuGlnGlyGlyTyrLeuGlyGlnIlyValAla 260
DB 978 ATGGGAGAGGAGAGTGTGGAGCTGCTCGAGGGGAGTACCTGGGGCAAAAGGTGGCC 1037
QY 261 ValIlyAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
DB 1038 GTGAAGAATATCAAGTGTATGTGACAGCCAGCCCTTCTGTGAGAGAGAGCCCTCATG 1097
QY 281 ThrIlyMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnIlyLeu 300
DB 1098 ACGAAGATGCAACAGAGAACTGGTGGCTCTCTGGGGGTATCTGACACAGGGGGCTG 1157
QY 301 TyrIleValMetGlnHisValSerIlyGlyAsnLeuValAsnPheLeuArgThrArgGly 320
DB 1158 TACATTGTCAITGAGACAGCTGAGCAAGGCAACCTGTGTAACCTTCTGCGGACCCGGGT 1217
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlyGlyMet 340
DB 1218 CGAGCCCTCGTAACACCGCTCAGCTCCTGCACTTCTCTGACGTGGCCAGGGCATG 1277
QY 341 GluTyrLeuGlySerIlyLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
DB 1278 GAGTACCTGGAGAGCAAGAGCTGTGCACCGACCTGGCGCCGCAACACTCTGGTGC 1337
QY 361 SerGluAspLeuValAlaIlyValSerAspPheGlyLeuAlaIlyAlaGlyArgGly 380
DB 1338 TCAGAGGAGCTGTGGCCAGGCTCAGCGACTTGGCTGGCCAAAGCCGAGCGAGGGG 1397
QY 381 LeuAspSerSerArgLeuProValIlySTPrThrAlaProGluAlaLeuLysHisGlyLys 400
DB 1398 CTAGACTCAAGCGGCTGCCCTCAAGTGGAGGGCGCCGAGGCTCTCAACACGGGAAG 1457
QY 401 PheThrSerIlySerAspValIlyPheArgPheGlyValLeuLeuTyrGluValPheSerTyr 420
DB 1458 TTCACCACCAAGTGGATGTCTGGAGTTTGGGGTGTGCTCTGGAGGCTCTTCATAT 1517
QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGlyLysGly 440
DB 1518 GGAGCGGCTCCGTACCTTAATGTCACCTGAAAGAGGTGTGCGAGGCGGTGGAGAGGG 1577
QY 441 TyrArgMetGluProGluGlyLysProGlyProValHisValLeuMetSerSerCys 460
DB 1578 TACCCCATGGAACCCCGAGGGGCTGTCAGGCCCGCGACGCTCATGAGCAGCTGC 1637
QY 461 TrpGluAlaGluProAlaArgProPheArgPheArgLysLeuAlaGlyLysLeuAlaArg 480
DB 1638 TGGAGGCGAGAGCCCGCCCGGCGGCTTCCGCAAACTGGCGAGAGCTGGCCCGG 1697
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
DB 1698 GAGTACAGCAGTGCAGGTGGCCCGACCTTCCTCTCAGGGCAGGAGCGCGAGCTCCACC 1757
QY 501 SerProArgSerGlnGluPro 507
DB 1758 TCGCCCGAGAGCCAGGAGCC 1778

```

RESULT 4
AAV44497

ID AAV44497 standard; cDNA; 1987 BP.

AC AAV44497;

DT 16-OCT-1998 (first entry)

DE Human matr cDNA.

KW Breast; cancer; matr; Csk homologous kinase; CHK; detection; diagnosis;
cytoplasmic protein; tyrosine kinase; ErbB-2; negative regulator;
mitogenic signalling; ss.

```

XX OS Homo sapiens.
XX FH Key
XX FT CDS Location/Qualifiers
XX FT 263..1846
XX FT /*tag= a
XX FT /product= CHK
XX FT /note= "CSK homologous kinase"
XX PN W09830704-A1.
XX PD 16-JUL-1998.
XX PF 07-JAN-1998; 98MO-US00420.
XX PR 16-JUN-1997; 97US-0876882.
XX PR 08-JAN-1997; 97US-0035228.
XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX PI Avraham H, Groopman JE;
XX DR WPI: 1998-399149/34.
XX DR P-PSDB: AAW64454.
XX PT Detecting breast cancer by detecting Csk homologous kinase
XX PT expression - especially in humans and use of Csk homologous kinase
XX PT in treatment or prophylaxis of breast cancer and for producing
XX PT medicaments
XX PS Disclosure: Fig 2: 54pp; English.
XX CC This sequence encodes a CSK homologous kinase (CHK) which is used in a
XX CC method of detecting cancer in breast tissue. The method allows diagnosis
XX CC of breast cancer in mammals, especially humans. It is based on the
XX CC discovery that a cytoplasmic protein tyrosine kinase, CHK, is expressed
XX CC in human breast tissue, but not in adjacent tissue. This protein can be
XX CC used to raise antibodies which can be included in compositions and
XX CC diagnostic kits for diagnosis of breast cancer. The presence of CHK in
XX CC breast tissue can also be determined using other standard methods (e.g.
XX CC Northern blotting) or by detecting nucleic acid sequences encoding all/a
XX CC portion of the protein (e.g. using hybridisation probes). Over-expression
XX CC of the receptor tyrosine kinase ErbB-2 has previously been associated
XX CC with the development of breast cancer, and CHK specifically interacts
XX CC with activated ErbB-2, and may function as a negative regulator of
XX CC ErbB-2 mediated mitogenic signalling. The compositions may also be used
XX CC to design drugs (e.g. which incorporate CHK analogues with greater
XX CC biological activity than CHK) and to identify CHK antagonists and
XX CC agonists for therapeutic use.
XX SQ Sequence 1987 BP; 369 A; 628 C; 672 G; 318 T; 0 other:
XX
XX Alignment Scores:
XX Pred. No.: 1,91e-192 Length: 1987
XX Score: 2591.50 Matches: 502
XX Percent Similarity: 99.21% Conservative: 1
XX Best Local Similarity: 99.01% Mismatches: 3
XX Query Match: 97.02% Indels: 3
XX DB: 19 Gaps: 1

```

US-09-977-260-2 (1-507) x AAV44497 (1-1987)

```

QY 1 MetaIaGlyArgGlySerLeuValSerThrPrArgAlaPheHisGlyCysAspSerAlaGlu 20
DB 263 ATGGCGGGGCGAGGCTCTCTGTTCTTCTGGCGGCAATTCAGGCTGTGATCTGTGAG 322
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaThrPheHisPropProValSerAla 40
DB 323 GAACCTCCCGGGGTGAGAGCCCGGCTCTCGACCTGGAGACCCCGCTCCGCTCAGCC 382
QY 41 ArgMetProThrArgArgTPrAlaProGlyThrGlnCysIleThrIlyCysGlyHisThr 60
DB 383 AGGATGCCACAGAGGCGCTGGGCCCGGCGACCCAGTGTATCACCAATGCGAGCACACC 442

```

QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
 DB 443 CGCCCCAAGCCAGGGAGCTGGCTTCCGCAAGGGCAGCGTGCACCATCTCGAGGCC 502
 QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGluGlyLeu 100
 DB 503 TGGCAAGACAGACCTGTACCGCGTACAGCACACAGTGGACAGAGGGGCTCTCG 562
 QY 101 AlaAlaGlyAlaLeuAlaArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 DB 563 GCACCTGGGGCGCTGGCGGAGAGGCCCTCTCCGACAGCCCAAGCTCAGCTTATG 622
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
 DB 623 CCGGCTTCCACAGGAGATCTCGGCGCAGAGAGCTGCCAGCAGCTGACGCTCCGAG 682
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal 160
 DB 683 GATGGCTGTCTCGTGGCGGAGTCCGCGCACCCCGCGCACTGCTGTGCTG 742
 QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
 DB 743 AGCTTTGGCCGCGACGTCATCCATCCGCTGCTGCACCGCGCACGCTCAGCATC 802
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspValGluHisTyrSerLysAspLys 200
 DB 803 GATGAGCGCTGTCTTCTCCACCTCATGAGATGTTGAGCATACACCAAGACAG 862
 QY 201 GAlaValIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerLysAla 220
 DB 863 GCGCCTATCTGCACCAAGCTGTGTAGACCAAGGGAACACGGGACCAAGTCCGCCAG 922
 QY 221 GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 DB 923 GAGAGCTGGCCAGAGCGGCGCTGTACTGACCTGACATTTGACATTTGGAGCAG 982
 QY 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyLysTyrLeuGlyGlnLysValAla 260
 DB 983 ATCGAAGGAGGAGTGTGAGCTGTCTCAGAGGTAGTACCTGGGGCAAAAGGTGGCC 1042
 QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet 280
 DB 1043 GTGCAAGATATCAAGTGTGATGTGACAGCCAGGCTTCTTGGAGGAGCGCGTATG 1102
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuGluGlyAlaIleLeuHisGlnGlyLeu 300
 DB 1103 ACGAAGATGCAACAGCAAGACGTGTGCTCTCTGGCGCTGATTCGACACAGGGGCTG 1162
 QY 301 TyrTleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 DB 1163 TACATTGTCTAGAGCAGCTGACAGCAAGGCGCACTGTGTAATTTCTGCGGAGCCGGGCT 1222
 QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet 340
 DB 1223 CGAGCCCTGTGAACACCGCTCAGCTCCCTGACAGTTTCTGTGACGCGGGCGAGGCGATG 1282
 QY 341 GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaIlaArgAsnIleLeuVal 360
 DB 1283 GAGTACTGTGAGAGCAAGAGCTGTGACCGGAGCTGGCCCGCAACATCTGTGTC 1342
 QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
 DB 1343 TCGAGAGACTGTGGCCAAAGTCACGACTTGTGGCTTGGCCAAAGCCGAGCGAAGGGG 1402
 QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 DB 1403 CTGAGACTCAAGCCGCTGCCCTCAAGTGAAGCGCGCCGAGGCTCTCAAAACAGGGG--- 1459
 QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
 DB 1460 TTACACAGCAAGTCGATGTCTGAGATTGGGGTCTGTCTCTGGAGAGTCTTCTCATAT 1519

QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 DB 1520 GAGCGGCTCCGTCACCTAAATGTCACTGAAAGAGGTGTGAGAGCGCTGAGAAAGGGG 1579
 QY 441 TyrArgMetGluProProGluGlyCysPheGlyProValHisValLeuMetSerSerCys 460
 DB 1580 TACCGCATGGAAACCCCGGAGGGCTGTCCAGGCCCCCTGCACGCTCATGACAGCTGC 1639
 QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
 DB 1640 TGGGAGGCACAGCCGCC-CCCGGCGCACCTTCCGCAAACTGGCCGAGAAAGCTGGCCCG 1698
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 DB 1699 GAGCTACGACAGTGTGAGGTGCCCGCCTCGTCTCAGGCGCAGAGACGCGGATCC-ACC 1757
 QY 501 SerProArgSerGlnGluPro 507
 DB 1758 TCGCCCGCAAGCAGAGGCC 1778
 RESULT 5
 ABR88791
 ID ABR88791 standard; cDNA; 1713 BP.
 AC ABR88791;
 AC
 DT 21-OCT-2002 (first entry)
 DT
 XX
 DE cDNA encoding human kinase related to tyrosine kinase family.
 XX
 KW Human; kinase; tyrosine kinase; kinase mediated disorder;
 KW hematopoietic disorder; cell signal transduction disorder;
 KW cancer; haemostatic; cytosolic; chromosome 19; gene; ss.
 OS Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 223..1470
 FT /*tag= a
 FT /product= "kinase"
 XX
 PN W0200252018-A2.
 XX
 PD 04-JUL-2002.
 XX
 PF 19-DEC-2001; 2001WO-US48546.
 XX
 PR 21-DEC-2000; 2000US-0741154.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Beasley EM, Shao W, Ketchum K, Di Francesco V;
 XX
 DR WPI: 2002-583568/62.
 DR P-PSDB: AA099575.
 XX
 PT New isolated human kinase proteins and genes, useful in developing
 PT drugs, as well as for diagnosing, preventing or treating disorders
 PT associated with defective cell signal transduction, e.g. cancer or
 PT haematopoietic disorders
 XX
 PS Claim 4; Fig 1A; 78bp; English.
 PS
 CC The present invention relates to the isolation of a novel human kinase
 CC related to the tyrosine kinase family, and polynucleotide sequences
 CC encoding it. The gene encoding the kinase of the invention maps to
 CC chromosome 19. The kinase may be used for identifying a modulator of
 CC the kinase, an agent that binds to the kinase, or for identifying
 CC other members of the family. The kinase may also be used to raise
 CC antibodies which may be used in immunoassays or drug screening assays.
 CC The sequences of the invention may be used for treating a disease or
 CC condition mediated by a human kinase such as haematopoietic disorders
 CC and other disorders associated with defective cell signal transduction

PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
genes -
XX
XX
PS Claim 2: Page 2592-2596; 2690pp; English.

CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (1) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (1). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (AB199202 to AB199912), encoding
CC the protein sequences in AB199202 to AB199912, or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator or when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
XX

Sequence 8212 BP; 1663 A; 2301 C; 2300 G; 1948 T; 0 other:

Alignment Scores:
Pred. No.: 9, 21e-110 Length: 8212
Score: 1540.50 Matches: 420
Percent Similarity: 34.26% Conservative: 23
Best Local Similarity: 32.48% Mismatches: 39
Query Match: 57.68% Indels: 816
DB: 24 Gaps: 12

US-09-977-260-2 (1-507) x AB199884 (1-8212)

QY 26 SerProArpPheLeuArGAlaTrpHisProProValSer-----AlaArg 41
Db 4184 AGCCCGAGTTTCTCTCAGAGTCTCTGCTCCCAATTCACACTTGCCTGTCTCA 4243
QY 42 MetPro-----ThArGArGTTPAlaProGlyThrGln 52
Db 4244 GTCCCTGACATACCTCTGCTCTGCTCTCCACAGCGCTGGGCCCTGGAGCTCAA 4303
QY 53 CysIleThrLysCysGluHisThrArGProLysProGluLeuAlaPheArgLysGly 72
Db 4304 TCATGACCAAGTGTGAGACTCTGCCCCAAGCCGGGTGAGCTTACCTTTGAAAGGCT 4363
QY 73 AspValValThrIleLeuGluAlaCysGluAsn----- 83
Db 4364 GACATGCTGACCATCTTGGAGGCTGTGAGT-GAGAGGGAGCTGAGAGGTAGTGTG 4422
QY 83 ----- 83
Db 4423 TGCCTTGGGGTGCCAGGGCCGACACCACAGCTCTTCCACATCTGTCCCTTCT 4482
QY 84 -----LysSerTrpTyrArGValLysHisHisThrSerGlyGlnGlyLeu 100
Db 4483 GTACAGACACAGACTGTGACCGAGCCAGCACCATGAGCGAGGAGAGGCTGTG 4542
QY 101 AlaIaGlyAlaLeuArGluArGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 4543 GGGCGCGTCTGTGGACAGCGGGAGGCCCTCTCCACAGACCCCAAGCTCAGCTCAG 4602
QY 121 Pro----- 121
Db 4603 CCGTAGTGGCACCTGGGGGGTGTGTGAGAGGCTCAAGCCCTTACCCCTGCTTAC 4662
QY 122 -----TrpPheHisGlyLysLysLysSerGlyGlnGly 131
Db 4663 AAGAGCCCTCTCCACACACATCCCTACAGATGTGTTTATGCAAGATTCGGGCGAGA 4722
QY 131 uAlaValGlnGlnLeuGlnProProGluLysPheLeuValArgLysSerAlaAr 151
Db 4723 AGCATTACAGAGCTGACAGCCACCGAGAGAGCGGCTTCTTGTGAGGGAATCAGCCG 4782

QY 151 gHisProGlyAspTyrValLeuCysValSerPheGlyArGAspValIleHisTyrArgVa 171
Db 4783 TCACCTGGAGACTATGCTCTGTGTGCTGATTTGGCCGTGACGTCATCCACTACCTGT 4842
QY 171 IleuHisArGAspGlyHisLeuThrIleAspValAlaValPhePheCysAsnLeuMet 191
Db 4843 TTTGCATTCAGATGGGACACTCACCATCATGAGGCGGTGTCTGTGAACCTGATGA 4902
QY 191 pheValGlu----- 194
Db 4903 CATGTGGA-GGTGACGTGTGGGCAAGGCTGGGAGAGGAGATGAAATATGATGTG 4961
QY 194 ----- 194
Db 4962 GGGCCAGGAACTGTCACTTCTTCTAGAACTGGGGTCACTCCCGAGGAGCTTAT 5021
QY 194 ----- 194
Db 5022 GTTCAACACCAACACCCCTCAAGTTCCATTTCACAAACCCCACTGTGGCAAAACC 5081
QY 194 ----- 194
Db 5082 GATCACCTTAGCCTGGCACTTAGAGAGAGAGTGTGGCAACTCACTGAACCTTGGGGTG 5141
QY 195 -----HisTyrSer 197
Db 5142 GATGTGGAGAGCGGTCTGTCAGAGGAGAGCTCATGCGCGGCTCCACACACTACAC 5201
QY 197 rLysAspLysGlyAlaIleCysThrLysLeuValArGProLysArGlyHisGlyThrLys 217
Db 5202 CAAGGACAAAGGGGCGCATCTGCACCAAGCTGTGAAGCCAAAGAGGAAACAGCGCA 5261
QY 217 sSerAlaGlnGlnGlnLeuAlaArg----- 225
Db 5262 GTCTGAGAGAGAGAGAGCTGCCAA-GGATATGAGAGCCATAGTGAGAAAGAGATCT 5320
QY 225 ----- 225
Db 5321 CGTGGCAGCAGGAGAGACTCCACCTCCAGACTCAGACCCCTAATTCATTGTCACCC 5380
QY 225 ----- 225
Db 5381 TCACACCCAGGCTTGGCTTGTAAAGCCCTGCGCTTCCAAACCTGTACACATTATCC 5440
QY 225 ----- 225
Db 5441 ACTGGGGCTATGATCATCAGAGCAAAATCGCAGTGGAGAGGTGCTGCTGCTTGG 5500
QY 225 ----- 225
Db 5501 AATCTTCAAGAGTGGCGGTAGGGTGGGGTGTAGAGTCTTTGGAAGGAGAGGGCTTTC 5560
QY 225 ----- 225
Db 5561 AAGGCAACTGCAGTCTGTGAGAGTACAGATTAAGAGACTATAGCCCTGAAAGATT 5620
QY 225 ----- 225
Db 5621 CCAGAACCTGTCAAACTCTCTGATTTCTGTGCTTGTCTGTGTGGTGTGAGCAAG 5680
QY 225 ----- 225
Db 5681 GGGAGGCCCTTCCAGAGGCAATGGGGGCGACAGGGGAGTCACTCTGCTCCTAGA 5740
QY 225 ----- 225
Db 5741 GGGAGGCTCTGAGCATGAGCTCAAGTATAGGAACCTAGGGGGGAGCTGTATTCTGATGA 5800
QY 225 ----- 225
Db 5801 AGGCTGGATGGCTGTGGCTTGCAGGGGCTGAATATAGTACAGAGAGTGGGTGCTTCA 5860
QY 225 ----- 225

Db 5661 TGCAGCCTTAGGCGAGCTGTGTGAGACAGGACATGCTCAGGTGTGCCAGGCTGTGGGAT 5920
Qy 225 -----
Db 5521 ATCAGCTCTCCGAGCTCTCAGTGAATGCAAACTGGGTGGGATGGCGAGAGGCTCTCTGC 5980
Qy 226 -----
Db 5981 CTCCTCTCAGTGGGCTCTTGTCTGTCTGACAGCTGGGCTACTGCACTGCACTGCACTCT 6040
Qy 235 uThleuclYalaglnileglylglulphleuYala-----
Db 6041 GACTCTGGGACACAGATTTGGAGAGGGAGTTGGAGG-TGAGAGGGGACACAGCAGG 6099
Qy 248 -----
Db 6100 CAGTGTTCAGAAATTGACCTTCAACTGAGAGTGACAGACAGGTGGCTGTGGGA 6159
Qy 248 -----
Db 6160 GTACGACACCCCTCAGCTCCAGGATCCTATTCAGAAAACCTTCCCTGTGTGTGT 6219
Qy 248 -----
Db 6220 TCTGGGCTTGGGTGACAGAGAGTAGACTTCAGGATGTCTCAGTGGCGCAGACCCGA 6279
Qy 248 -----
Db 6280 AATTATGCAAAATGGCGAGACCTGTGGTTCCTTCATGAGAACAGAAATTTGATTCTG 6339
Qy 248 -----
Db 6340 AATGAGACGCCATCGATTGTCATGATTAATTTTTCCTCATTCAGGCGCTGGGG 6399
Qy 248 -----
Db 6400 GATGAGACGAGCCTCTGATGATTTTCAACTCTGTGCCAGGCGACACCCCGAGCC 6459
Qy 248 -----
Db 6460 CTCAGTGGGGATTCCAGGAGGGGTGTGACAGGCTGTCTGTCTGTCTGTCTGTCTGTG 6519
Qy 249 -----
Db 6520 TGGGCGCTGACCTCTGTCTCTCTCCACAGCGCTCTTACAGGGGTGAGTCTGGACAGAA 6579
Qy 258 sValAlaValLysAnlleLysCysAspValThraGlnAlaPheLeuAspGluThrAl 278
Db 6580 GGTGGCTGTGAAGATATCAAGTGTGATGACAGCCAGGCGCTCTGATGAGACGCG 6639
Qy 278 aValMethr-----
Db 6640 TGTATGACGTGAGTGTGGGGTGGGCGCAGGCTCCAGGCTGAGGCTCCAGGCTGTATAC 6699
Qy 282 -----
Db 6700 TCACCTATGTCCTCCCTCTGATCCCTGACACCCAGAGACCTGAGAGCAGAGACCTAG 6759
Qy 289 aLaGleuLeuGlyValIleLeuHsGlnGlyLeuTyrIleValMetGlnHsValSerL 309
Db 6760 TGCACCTCTGGGTGTGATCTGTGACACGCGCTGTACATGTGAGAGCAGTGAGCA 6819
Qy 309 ys-----
Db 6820 AGGTGTGACAGGGGTGGGTGGGTGGGACAGAAAGGTGGCCCAACAGCTTGAGAGACCC 6879
Qy 310 -----
Db 6880 CAAGTCCCTCCCTTAACAGAGGCACTGTGAACTCTGTGCGAGCGGGCGCTGTCT 6939
Qy 323 LeuValAsnThrAlaGlnLeuLeuGlnPhe----- 332

Db 6940 CTGTGACACCTCTCAGCTTCTGAGATTTGCTGTGTAAGTACATTCATCCAGAGAGT 6999
Qy 333 -----
Db 7000 AGTCCCAAGTGGAAAGGAACTAGGCGCTGACTCCAAAGTGAAGCGGCGCTGTGTCT 7059
Qy 334 Leu-HisValAlaGlnGlyMetGluTyrLeuGluSerLysLysLeuValHisArgAspL 353
Db 7060 CTAGTCAGTGTCTTAAGAGCATGATACCTGAGAGAGCAAGAGCTGTGACACCGGAGC 7119
Qy 353 euaAlaAlaArgAsnIleLeuValSerGlnAspLeuValAlaLysValSerAspPheGlyL 373
Db 7120 TGGCTGTGGGAACATCTGCTGTGAGAGCTGTGGCCAAAGCTCAGTCACTTGGCT 7179
Qy 373 euaAlaAlaGlnuArgLysGlyLeuAspSerArgLeuProValLysThrPrrAlaP 393
Db 7180 TAGCCAGGACAGAGCGCAAGGGGTGAGTCAAGCGGCTCCAGTCAAGTGAAGAGGAC 7239
Qy 393 roGluAlaLeuLysHisGly-----
Db 7240 CTGAGCTCTCAAAAACGGGGTGAAGACCTTCTACCTCTCTGTGGATTATGGGTTCC 7299
Qy 399 -----
Db 7300 TCAGAGGCGCTGACGTTCTTGATTTGCTGTGACATATGACACCCCTCAACCTC 7359
Qy 399 -----
Db 7360 TCTGAGCTCAAGGGGAAATGGTTCCTTCACAGTGAAGCCCTAGACAACTTCCTC 7419
Qy 399 -----
Db 7420 CTAGACTTCCCTCAGCTGANTCCCATCTCTGAAGAGGCGCTGGCTTGAAGACAGAG 7479
Qy 399 -----
Db 7480 AGGAGTAGCCTGGGATTTTCCAGGAGGCCAGAGGGGCTCTCACCTGACCTC 7539
Qy 400 -----
Db 7540 TGCCTCAACAGCGGTTCTCCAGCAAGTGTGAGTGTGGATTTGGGTGTGTGTGG 7599
Qy 416 luValPheSerTyrIleArgAlaProTyrProLysMetSer----- 429
Db 7600 AAGTCTTCTTATGAGAGAGCCCATACCCCAAGATGTG-TGAGCAGACCTTAGGAGG 7658
Qy 429 -----
Db 7659 GAGGCGAAGACAGACATTCAAGCAGCTGTGTGACATGTCTCCACATGAACCTCT 7718
Qy 429 -----
Db 7719 GAGCAGCCCTGAGTGTGATGTCTGTATGACAGCTGTGTATCTGTCTCACTGACTGCC 7778
Qy 430 -----
Db 7779 CTCCAAGCCAGGCTGACACCCCTGTGTCCACAGATGCTTAAGAGGTTTCAAGAGCTGTGG 7838
Qy 438 luLysGlyTyrArgMetLupProPogluGlyCysPProGlyProValHisValLeuMetS 458
Db 7839 AGAAGGTTTACCGCATGTGAGGCCCGGATGGCTGCCAGGCTGTGTGACACACCTTCAGG 7898
Qy 458 erSerCysTrpGlnuLagLupProAlaArgArgProPheArgLysLeuAlaGluLysL 478
Db 7899 GTAGCTGTGGAGGACAGGCTGGGCGCGACACCTTCCGAAATATGAGGGAAGC 7958
Qy 478 euaAlaArgLupLeuArgSerAlaGlyAlaProAlaSerValSerLysLinspAlaAsp 498
Db 7959 TGGGCGTGAAGCTCGCAGTGTGGGTGTCTCCGCGCCCGCTGAGGAGACAGAGGTGAGG 8018
Qy 498 lySerThrSerProArgSerGlnLupPro 507
Db 8019 GCTCAGCTCCACAGAGGAGCCAGGACCC 8047

RESULT 7	FT	Intron	13036..13941	/tag= s	variation
ABK88792	FT	exon	/number= 9		
ABK88792 standard; DNA; 16389 BP.	FT		13942..14178		
AC	FT		/tag= t		
XX	FT	variation	/number= 10		
XX	FT		(replace 1130, A)		
DT	FT		/tag= u		
21-OCT-2002 (first entry)	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4245, A)		
	FT		/tag= v		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4248, T)		
	FT		/tag= w		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4250, A)		
	FT		/tag= x		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4255, A)		
	FT		/tag= y		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4258, T)		
	FT		/tag= z		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4260..4261, T)		
	FT		/tag= aa		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4266..4267, T)		
	FT		/tag= ab		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4269..4270, T)		
	FT		/tag= ac		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4273, T)		
	FT		/tag= ad		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4275, A)		
	FT		/tag= ae		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4401, T)		
	FT		/tag= af		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4551..4552, C)		
	FT		/tag= ag		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 4739, G)		
	FT		/tag= ah		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 6309, A)		
	FT		/tag= ai		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 8753..8754, T)		
	FT		/tag= aj		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 10797, T)		
	FT		/tag= ak		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 10806, G)		
	FT		/tag= al		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 11199, T)		
	FT		/tag= am		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 11201, T)		
	FT		/tag= an		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 11319, C)		
	FT		/tag= ao		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 12238, C)		
	FT		/tag= ap		
	FT	variation	/standard_name= "Single nucleotide polymorphism (SNP)"		
	FT		(replace 14739, T)		

```
FT FT /*tag= aq
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT FT (replace 15375, C)
FT FT /*tag= ar
FT FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX PN MO200252018-A2.
XX PD 04-JUL-2002.
XX PF 19-DEC-2001; 2001WO-US48546.
XX PR 21-DEC-2000; 2000US-0741154.
XX (PEKE ) PE CORP NY.
XX PI Beasley EM, Shao W, Ketchum K, Di Francesco V;
XX WPI: 2002-583568/62.
XX P-PSDB; AAU99575.
XX DR
XX New isolated human kinase proteins and genes, useful in developing
XX PT drugs, as well as for diagnosing, preventing or treating disorders
XX PT associated with defective cell signal transduction, e.g. cancer or
XX PT hematopoietic disorders
XX PS
XX PS Claim 4; Fig 3A-F; 78pp; English.
XX CC The present invention relates to the isolation of a novel human kinase
XX CC related to the tyrosine kinase family, and polynucleotide sequences
XX CC encoding it. The gene encoding the kinase of the invention maps to
XX CC chromosome 19. The kinase may be used for identifying a modulator of
XX CC the kinase, an agent that binds to the kinase, or for identifying
XX CC other members of the family. The kinase may also be used to raise
XX CC antibodies which may be used in immunoassays or drug screening assays.
XX CC The sequences of the invention may be used for treating a disease or
XX CC condition mediated by a human kinase such as hematopoietic disorders
XX CC and other disorders associated with defective cell signal transduction
XX CC (e.g. cancer). The polynucleotide sequences encoding the kinase are
XX CC useful for isolating and purifying the kinase, and as probes or primers.
XX CC The present sequence encodes the human kinase of the invention.
XX SQ
SQ Sequence 16389 BP; 3511 A; 4901 C; 4515 G; 3462 T; 0 other;
Alignment Scores:
Pred. No.: 7.56e-104 Length: 16389
Score: 1469.50 Matches: 501
Percent Similarity: 21.59% Conservative: 0
Best Local Similarity: 21.59% Mismatches: 6
Query Match: 55.02% Indels: 1819
DB: 24 Gaps: 12
US-09-977-260-2 (1-507) x ABK88792 (1-16389)
QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
Db 7229 ATGCGCGGGGAGGAGCTCTGTGTCTCGGCGGCATTTCAGCGCTGATTTGCTGAG 7288
QY 21 Glu----- 21
Db 7289 GAACTTCCCGGGGTAAGATCACTTCCCAAGGGGCTTGGGAGCCAGTTCTGGGTCCC 7348
QY 21 ----- 21
Db 7349 ATCCAGATGCTCTGGGAGAGCTCTGTCTGCTGCTCCAGCATGCGCCCACTGACT 7408
QY 21 ----- 21
Db 7409 GGGCTGAAGTGTAGGAACGTGACCGTGGTCCAGCCAGCTCCACACTGACTCTGT 7468
QY 22 ----LeuProArgValSerProArgPheLeuAlaArgAlaPheHisProProArgValSerAla 40
Db 7469 CCGCTCCCAAGGTG-AGCCCGCGCTTCTCCGAGCGTGGCACCCCGCTCCCGTCTGACGC 7527
```

```
QY 41 ArgMetProThr----- 44
Db 7528 AGGATGCCACGAGGTAGTGTCTAGACCTGCTTCTGCTCCCGGGTCTTCTTCAAC 7587
QY 44 ----- 44
Db 7588 TCCGTCCACACCCGAGAGCCCTTCTGTAAGGGCTGATCTGTGCCCCCTACCCCTTC 7647
QY 44 ----- 44
Db 7648 TGTGCGCTCGGCTGAGCTTTCACATCATTTGGCTTCTGACACTTGTGCCAGGC 7707
QY 44 ----- 44
Db 7708 GCGTCTTGTGTACTAGATTTCTCTGTAGACACCCACATCCCTTCTTCCCCC 7767
QY 44 ----- 44
Db 7768 GATGTCTGCGTGTCTTCTGTGCTCTCCGCTCTTTCATCTATACCTTCTTCCC 7827
QY 44 ----- 44
Db 7828 TCTGTCTCTTTCCTCCCTCCGCTGCTCTGCTCCACATCTCCGTGATCTCTCCCTT 7887
QY 45 -----ArgArgTrpAlaProGlyThrGlyHisIleThrIly 56
Db 7888 GCCCTCTCCCGGCGCCCTCTGCAGAGGCGCTGGGCCCGGACCGATATACCAA 7947
QY 56 sCyGluHisThrArgProLysProGlyGluLeuAlaPheArgIleValAlaIle 76
Db 7948 ATGGAGACACCGCCCGCAAGGACGAGGAGCTGCGCTTCCGAAGGGCGAGCTGTCAC 8007
QY 76 rIleLeuGluAlaGlyGlu----- 82
Db 8008 CATCTGAGGCTGCGA-GGTGAGAGTGGCGCGGCTGTGGTGTGGGCTGGGCG 8066
QY 83 -----As 83
Db 8067 GCTCCACGAGACACCCCGCCACACCCCGCCACACTAACCCCTTCTCTCCCTGGCAG 8126
QY 83 nLysSerTrpTrpArgValLysHisThrSerGlyGluGluLeuAlaAlaGlu 103
Db 8127 CAAGAGCTGGTACGCGTCAAGACACACACAGTGAAGAGGGGCTGTGGCAGCTGG 8186
QY 103 yAlaLeuArgLysArgLysAlaLeuSerAlaAspProLysLeuSerLeuMetPro----- 121
Db 8187 GCGCTCGGGAGCGGAGGCCCTCTCCGACAGCCCAAGCTCAGCTCAGCTGAGT 8246
QY 121 ----- 121
Db 8247 GGGGAGACAGGGGCTGGGTAGGGGACAGCAAGTACCCCTCCACAGCCAGCT 8306
QY 122 -----TrpPheHisGlyLysIleSerGlyGluAlaAla 133
Db 8307 GACCACCCCTTCCGTGGCGCCAGAGTGTTCACAGGGAATCTCGGGCGAGAGGCT 8366
QY 133 lGlnGluLeuGlnProProGluAspGlyLeuPheLeuValArgLysAlaArgHisPr 153
Db 8367 CCAGCAGCTGAGGCTCCCGAGATGGCTGTCTGTGGGAGTCCGGGCCACACC 8426
QY 153 oGlyAspTrpValLeuGlyValSerPheGlyArgAspValIleHisTrpArgValLeu 173
Db 8427 CGGCGACTACGCTCTGTGCGGTGAGACTTTGGCGGAGCTCATTCACCTACCGCTGCTCA 8486
QY 173 sArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVa 193
Db 8487 CCGGAGGCGCACCTCAATCATGTAGGCGGTCTTCTTGCACACCTCATGACATGTGT 8546
QY 193 lGlu----- 194
Db 8547 GGAAGTCTGTGACCCAGAGGCCCGCCCGCTCTCCGTGAGTGGGCGAGGCCAGGAGCTCC 8606
```

QY 194 ----- 194
Db 8607 GGAGACTCCCTACGTAGAGATAGAGGGGGGTACCTGACCTCCATCCTTCCAGC 8666
QY 194 ----- 194
Db 8667 AGCTGGGCGACACACATCCCTACCTCTGTGTGATGACCCGGGACCCTGAGCTGTGT 8726
QY 194 ----- 194
Db 8727 GCTTCCCAATGTGAGATGAAAGGGGTATTCCCAAGGGCCCCCAAGGACCCCAAGGC 8786
QY 194 ----- 194
Db 8787 CCCATCCCATCTCAGATGGGCGACAGAGAGACCCCAAGTAGATTCCCTTAATGCA 8846
QY 194 ----- 194
Db 8847 GAAATAGGGGGGTGATCCCTATGAGACTGCCCAATCTACAGTGGCTGAGTCAGCCTGG 8906
QY 194 ----- 194
Db 8907 TTCTACCTCAGGAGACTAGAGTACTCTGTGGGACCCCCAGAGACCCCAAGTTGAT 8966
QY 194 ----- 194
Db 8967 ACCAATCAGAGAGACTCCCCCTCTTTGGCTTGCCCTTCCGTTCTATGAAACC 9026
QY 194 ----- 194
Db 9027 AGCCTTCTCTCTCTCCCGGTCCACCCACCACCCAGAGGCCACCCGGGAGCGAAC 9086
QY 195 -----H1 195
Db 9087 AGGATGCTGGGGTTCCCTCCCTGGGGCTGGGCTCATGGCTGCCACCATCCTGCAGCA 9146
QY 195 -----H1 195
Db 195 sTyrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLysHisG1 215
Db 9147 TTACACCAAGAGCAAGGGGCTATCTGCACAGCTGTGTAGACCAAGCGGAACACGG 9206
QY 215 yThrLysSerAlaGluGluGluLeuAlaArgAla----- 226
Db 9207 GACCAAGTGGCGCGAGGAGAGAGCTGGCCAGGGG-TagGGGAGGCCCAAGAGGCGAAGACC 9265
QY 226 ----- 226
Db 9266 CATTCCCTACCCAGCTTACCCAGTCCGGGAAGAGAGGCCCTGGGGCCCCGACCTTG 9325
QY 226 ----- 226
Db 9326 AGGCCAGATCACCCAAAGCTGTCTCATGCTCCAGCTATGGGATAGAAAGCTGGGCTGCC 9385
QY 226 ----- 226
Db 9386 TGGGGGTGGCCACAGACACAGATCAGTCTTATTATCATAGTCCCGACCTTCCGCACT 9445
QY 226 ----- 226
Db 9446 CATCTGTACCATCATCACCATCTGAGCCAGAAACCTGGCCTGAACCTCCACCCAGCC 9505
QY 226 ----- 226
Db 9506 CATCTCAGACATTCCTCTCTGGCCAGGTGCCAAGACCTGGTGAATCTTCTCCCTCC 9565
QY 226 ----- 226
Db 9566 CCTCCCAAGCCAGCCCTCAGGCCCTCCCATCAGACCCGCTTCTTCCAGAGGGGAGAT 9625
QY 226 ----- 226
Db 9626 CCCCTTCCAGCCAGGCTGTGATGACAGCTCAAGGTTATCTTCTATCATGAACAAT 9685
QY 226 ----- 226

Db 9686 GCCTCCAGGAGAGAGGCCCAAGAGCTCCACACACTTCCAGAACTGCCTGATGGCTG 9745
QY 226 ----- 226
Db 9746 AGCCACCTTCCAGCCTCATCTGCAAAAGCCCTTCTCCACACATCCACCCCTCCAGGT 9805
QY 226 ----- 226
Db 9806 CAACGTGATCATGTTTTCTTCATGCTCTGAGCCATTGTATATTCTGTTCCCTTACCC 9865
QY 226 ----- 226
Db 9866 TGAATGCTTTTCCAGCCCTTAAGTAGTGAATCTCAATATTTAAGTCAAAAGCAAT 9925
QY 226 ----- 226
Db 9926 GCCTCCTGACCCCATGTCTGTCTCTTGAAGTATCTCTGCTGGGAACTGTAC 9985
QY 226 ----- 226
Db 9986 CCATCAGGATGSCATCCAGGTGAGACTAGCATCACCTGGCCAATGATCTTAAAGAC 10045
QY 226 ----- 226
Db 10046 CGTGGCATTTGGAGGGGCACTTTCTGAGAGGAAAGAACGCTGCAGAACCGACAG 10105
QY 226 ----- 226
Db 10106 AGGGCAGTGGAGGGGACACACAGAGAGTCTTGGAGAAAGGTCAGAGGAGCTAGAGCCC 10165
QY 226 ----- 226
Db 10166 TGAATATTCTGSAACCTTATGTGGTCTCATGTGTGTTTTTGTCTGTGTGTTT 10225
QY 226 ----- 226
Db 10226 CTGGGGTGTGACAGGTAAAGAACTGTAGGAGAGAGCCAGAGGATTTGGAGGAGAC 10285
QY 226 ----- 226
Db 10286 AAGAGACCCAAAGAGAAAGGCGAGCTTATCTGTGAGAACTGAAGTGTTCCTGGCAC 10345
QY 226 ----- 226
Db 10346 TACCCCTTGGGCTGACTGGAGTCTCTGGAGGGCAAGTGGGGTCTGAGGGTCTACGC 10405
QY 226 ----- 226
Db 10406 CAGGACAGCCCGCAGCTTGTACTATTACACACACTATCCCTCTGAAGCCCTTTTGT 10465
QY 226 ----- 226
Db 10466 TCAAGTCAACCTGAGATAGTCTGTGTTCGTTGCAACAACAATATTTGTACAGAGATGGG 10525
QY 226 ----- 226
Db 10526 GTGTCTGTATATGTCTAGGGGCTATGCCAGGCCAGAGCCACTCAATATGTATCCA 10585
QY 226 ----- 226
Db 10586 CAGCTGCAGCACTGTGAGAGACATCACTAGTCTCTCACCAAGTAGAAGCCCAATGGGA 10645
QY 227 -----GlyTrpLeu 230
Db 10646 GGATTTAGGGGGCCATTACCCCTGACTGGGTCTTCACTCCACAGCGGGCTGTACT 10705
QY 230 uAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGluPheGlyAla----- 248
Db 10706 GAACCTGCAGCAATTTGACATTGGAGCACAGATCGGAGAGAGATTTGGAGG-TCAGCT 10764
QY 248 ----- 248

Db 10765 GCGGACGTGGTGGGGGTGCTGAAGAAGATGGAGGTGCATATTGAGCTTTGGAGTC 10824
QY 248 ----- 248
Db 10835 ACAGAGCTGAATTCAAAACCCACTTAAACCACCTACTAGTGTGCACTGAAGACGTGCA 10884
QY 248 ----- 248
Db 10885 AATCCTGTAGCCTCAGTTTCCCTCCCTGGAACAGGGGGTATGGCAGCTGCATGAGGA 10944
QY 248 ----- 248
Db 10945 GTAAGTGAATCAACAGCGGCTCAGTCAACCTTAGTACTTCTTACACTGCGAGGTC 11004
QY 248 ----- 248
Db 11005 TTGTATGCTAGTCCCACTGTTTGGAAATGTTTGAACACCAATCCCTGAATTAATTATTC 11064
QY 248 ----- 248
Db 11065 TGGAGGCTGTGGGTTTAGATCAGGGGTGCGGACACTCTCTGTAAGGCCAGATTGTA 11124
QY 248 ----- 248
Db 11125 AATATTTCTGCTTTGTGGCCAGGTAGCTGCTGCACCACTACTGAATCCTGCCCTTGT 11184
QY 248 ----- 248
Db 11185 GGAGTGAAGCAGCACAGACATATGTAAACATTTGGGCGTGCCCATGAATAACTATGAC 11244
QY 248 ----- 248
Db 11245 CCAGAAATTTGAATTTTGTATATGTCATCATGTCTGTGAACAGACATCTTTGATT 11304
QY 248 ----- 248
Db 11305 TGTTTTACACCAATTAAACATGCAAAAACAGCAGCGGTGCTCATGCTGGAA 11364
QY 248 ----- 248
Db 11365 TCCACGACCTTGTAGTGGGGGCGAGAGATCCCTTGAGCTCAGAGATTGAGTGCAG 11424
QY 248 ----- 248
Db 11425 CCTGACAAAGTAGGAGAGACCCCATCTGTACAAAAAAGCAGGCAAGTGGAGCAT 11484
QY 248 ----- 248
Db 11485 GCCTGTAGTCCAGCTACTCAGAGAGCGGAGTGGAGAGATCCCTTGAACCCAGGAATTT 11544
QY 248 ----- 248
Db 11545 GAACCTGCAGTAGCTGTATCAGCCCACTGCACCTTACCTGGGTGAGAGATGAGACCC 11604
QY 248 ----- 248
Db 11605 TGTATCAATATTTTAAAAATTAATAATGTGGCCGGGACAGTGGCTCATGCTGTATGCC 11664
QY 248 ----- 248
Db 11665 CAGCACTTGGAGGCTGAGCGGCGGAGATCAGAGAGTCATGATCGAGACCATCTGTGC 11724
QY 248 ----- 248
Db 11725 TTAACATGTTGAACCCCATCTCTACTAAAAAATTTAGCCGGCGGTGG 11784
QY 248 ----- 248
Db 11785 TGGTGGGCGCTGTAGTCCACACTTGGGAGGCTGAGGCAAGGAATGGGCTGACCTG 11844
QY 248 ----- 248
Db 11845 GGAGCGAGACTTGCAGTAGAGAGATGGCGCACACTGCATCCAGCCTGGCGGACAGAG 11904

QY 248 ----- 248
Db 11905 CAAGACTCGCTCAATAATTAATTAATTAATAATAATAATAATAATAATAATAATTTT 11964
QY 248 ----- 248
Db 11965 GAGCTGAAGGCTTTGCANAAAGAGGTATAGTATGATTTGGCCCCGAGGGGGTGGGCT 12024
QY 248 ----- 248
Db 12025 TCACAGACCCCTGGCTTAGATGTTCTCTTTGTTTTGTTTTGTTTTGTTTTGTTTTT 12084
QY 248 ----- 248
Db 12085 GAGATGAGTTTCACTCTGTCTACCCAGGCTGAGTGCATGCGACAAATCTTGGCTCAGTG 12144
QY 248 ----- 248
Db 12145 CAACCTGTGCTCCGGGTTCAAGTATTTCTCGCTCAGCTCCCGAGTAGTGGAT 12204
QY 248 ----- 248
Db 12205 TACAGGCAATGCGCACACACCTGTAGTTTGTATTTTCAATTAGAGACAGGTTTCTC 12264
QY 248 ----- 248
Db 12265 CATGTTAGTCGGGCTGTCTTCACTCCTGACCTCAGTGAATCCGCTGCTCAGCCTCC 12324
QY 248 ----- 248
Db 12325 CAAGTGTGGGATTACAGCGGTGAGCCACCGCCAGAGCCAGATGTTCTGTATTT 12384
QY 248 ----- 248
Db 12385 GGGTGTGAGCTTGCAAGGGCGGAGGGGTGAGAGACATGGGCTGTGTTCAAGTCT 12444
QY 248 ----- 248
Db 12445 GGGGACCCGACCTGTGTAAATTTACATGGGACCGGCTGTGAGCGGCTGCTGCTG 12504
QY 248 ----- 248
Db 12505 TCTCTCTGCTCCCTGTTTGTGGGGTCTGATCTGACCCCGTTCACCCACCCACCCGCC 12564
QY 249 ----- 267
Db 12565 ACCTGCTCTGAGGGGTGAGTACCTGGGCAAAAGGTGGCGGTGAAGATATCAAGTGA 12624
QY 267 pValThrAlaGlnAlaPheLeuAspGluThrAlaValMetThr----- 281
Db 12625 TGTGACAGCCAGGCTTCTCTGAGACGAGCGCGCTCATGAC-GTGAGTCCAGGTGG 12683
QY 281 ----- 281
Db 12684 GCTGGGACCTGTGGAGCGGGGGGTCCAGCCCTGCTCAGCCACCCACCCAGCCGCC 12743
QY 282 ----- 300
Db 12744 CAGGAAGATGCAACAGAGAACCTGGTGGCTCTCGGGCGTATCTCACACAGGGCT 12803
QY 300 uTyrIleValMetGluHisValSerLys----- 309
Db 12804 GTTCAATTGTATGAGACAGCTGAGCAAGGTGGGGGGGGCCCAAGCGGGGAGGGGCCCA 12863
QY 310 ----- 313
Db 12864 GCGAGGAGAGACCCCAACATCCCGGCGCATCTCCACCCCAAGGCAAGCTGCT 12923
QY 313 IAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGlnPheSe 333
Db 12924 GAACCTTCTGCGGACCGGGGCTGAGCCCTCTGAACACCCGCTCAGCTCCTCAGATTTC 12983

```

QY 333 rLeu----- 334
Db 12984 TCTGTAAAGTGGAGACTCTCAAGGAGTGTGCGGACTGGGGTGTGCGGGCGAGATTCAG 13043
QY 334 ----- 334
Db 13044 ATCCATCAGAGGGAACCTGAGGACGCGGGGAGGAGGCTTTGTGAGCTGTGTGCCAGG 13103
QY 334 ----- 334
Db 13104 CACCCGAGCTTCTGAGCCCTGATGTGGGATCTCTGACCTGAGCCGCCACTGCCCC 13163
QY 335 -----HisValAlaGluGlyMetGluTyrLeuGluSerLysLysLeuValH 350
Db 13164 CTACTACCCCGACGCTGGCGGAGGATGAGTACCTGAGAGCAAGACAGCTTGTGC 13223
QY 350 IsArgAspLeuAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSera 370
Db 13224 ACCGCACTGGCGCGCCCGCACATCTGTGTCAGAGGACCTGTGGCCAGAGTACGCG 13283
QY 370 spPheGlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysT 390
Db 13284 ACTTGGCTGGCCCAAGCGAGCGAGGAGGCTAGACTCAAGCCGCTGCCGTCAAGT 13343
QY 390 rPThrAlaProGluAlaLeuLysHisGly----- 399
Db 13344 GAGCGGCGCGCGAGGCTCTCAACACGCGGTGAGCCCTGCCCTCACATACCCCTGGGGCT 13403
QY 399 ----- 399
Db 13404 TTGGGTCGCCCGAGCTGTGCTGTATGACCCTGGGTATGTCTCTTGGCTCTGTAAACT 13463
QY 399 ----- 399
Db 13464 CCAAGGCAATGGCTATGCTCTCCCAAGAACTCTTGCCCATATTTGTTCTCGTAGCC 13523
QY 399 ----- 399
Db 13524 CCCCTGTGGGCTCAGTTTCTCCAGCTGTGMAAAGGCGTGGCTCAGACTGAAAGGAA 13583
QY 399 ----- 399
Db 13584 GAATACCAAGGCTCTTCATGCATTAAGGCTAGCTAGACTCACTGCGGGCGGCTTG 13643
QY 399 ----- 399
Db 13644 GGGGAGACTCAATTAGAGAGAACAGAGAGTGAAGGTCAACCAAGAGGCTTCTGTG 13703
QY 399 ----- 399
Db 13704 AGGAGCAGGGGCTGAGCAGAGAAACCTGAGGGGCTCTCTCCACCCCGCTCGGGCC 13763
QY 400 -----LysPheThrSerLysSerAspValTyrSerPheGlyValLeuLeuTyrGluValP 418
Db 13764 CACAGAAAGTTCCACACAGAGTGGAGTGTCTGGAGTTTGGGGTGTCTGGAGAGTCT 13823
QY 418 heSerTyrGlyArgAlaProTyrProLysMetSer----- 429
Db 13824 TCTCATATGAGCGGGCTCGCTACCTAAATGTT-GAGCGGGGCTCCAGAGGAGCACTG 13882
QY 429 ----- 429
Db 13883 GGTTCGGGCGAGTCCAGAGGCTGTGCGCTGACCCCTGCGCCAGCGCTGTCACAGT 13942
QY 430 --LeuTyrGluValSerGluAlaValGluLysGlyTyrArgMetGluProGluGlyC 449
Db 13943 CACTGAAGAAGGTGTGCGAGCGCTGAGAGGGGTACCGCATGGAACCCCGGAGGCT 14002
QY 449 ysProGlyProValHisValLeuMetSerSerCysTyrPheGluAlaGluProAlaArgP 469
Db 14003 GTCCAGGCGCCCGTGCAGCTCTCATGAGCAGTGTGAGGAGAGAGCCGCGCGCGG 14062
QY 469 roProPheArgLysLeuAlaGluLysLeuAlaArgGluLeuArgSerAlaGlyAlaProH 489

```

```

Db 14063 CACCCCTCCGCAAACTGCGCGAGAGACTGCGCCGAGGAGCTACCCAGTGCAGTGCAG 14122
QY 489 IsSerValSerGlyValAspAlaAspGlySerThrSerProArgSerGluPro 507
Db 14123 CCTCCGCTCAGGGGACAGAGCGCCGAGGCTCCACTGTGCCCCGAGCCAGAGAGCC 14178

RESULT 8
AAZ29701
ID AAZ29701 standard; cDNA; 2187 BP.
AC AAZ29701;
XX 22-MAR-2000 (first entry)
DE Wild-type human c-Src tyrosine kinase cDNA.
KW Angiogenesis; tyrosine kinase; Src; inhibition; activation; modulate;
KW human; viral expression vector; replication competent; mutant Src;
KW inflammatory disease; arthritis; rheumatoid arthritis; restenosis;
KW diabetic retinopathy; osteoporosis; cancer; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 134..1486
FT /tag- a
FT /product- "Human c-Src tyrosine kinase"
FT /note- "Src used to modulate angiogenesis"
XX
XX WO961590-A1.
XX 02-DEC-1999.
XX
XX 28-MAY-1999; 99MO-US11780.
XX
XX 29-MAY-1998; 98US-0087220.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX Cheresch DA, Ellicelri B, Schwartzberg PL;
XX WPI; 2000-116335/10.
XX P-P-SDB; AAY44448.
XX
XX Using tyrosine kinase Src for modulating angiogenesis in tissues useful
XX in, e.g. treatment of chronic articular rheumatism -
XX
XX Claim 1; Fig 3; 80pp; English.
XX
XX The present sequence is the cDNA, encoding the wild-type human c-Src
XX tyrosine kinase. This sequence encoding the Src protein can be used to
XX modulate angiogenesis. When the Src protein is inactivated, angiogenesis
XX is inhibited, while when it is activated, angiogenesis is potentiated.
XX The modified or variant Src can be used to treat inflammatory diseases
XX like, arthritis, rheumatoid arthritis, diabetic retinopathy, restenosis,
XX osteoporosis and cancer associated disorders.
XX
XX Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other;

Alignment Scores:
Pred. No.: 1.16e-87 Length: 2187
Score: 1247.50 Matches: 238
Percent Similarity: 71.59% Conservative: 82
Best Local Similarity: 53.24% Mismatches: 118
Query Match: 46.71% Indels: 9
DB: 21 Gaps: 3

US-09-977-260-2 (1-507) x AAZ29701 (1-2187)
QY 40 AlaArgMetProThrArg-----ArgThrAlaProGlyThrGlnCys 53
Db 116 GCCAGAGCTCTCGAGAGATGTCAAGCAATACAGGCGCGCTGCGCATCCGCTACAGAAATGT 175

```

QY 116 ThrLysCysGlnHisThrArgProLysProGlyLeuLeuAlaPheArgLysGlyAsp 73
 176 ATGGCCAGATACAACTCCAGCGACATCCGAGAGAGACCTGCCCTTCGCAAAAGAGAC 235
 QY 74 ValValThrIleLeuGlnAlaCysGlnAsnLysSerTrpTyrArgValLysHisHisThr 93
 236 GTCTCAACCAATTGTGGCCCTCACCAGAGAGCCCACTGATCAAAAGCCAAACCAAGGTG 295
 QY 94 SerGlyGlnGlyLeuLeuAlaIleGlyAlaLeuArgLysGlnValAlaLeuSerAla 113
 296 ---GGCCGTAGGGACATATCCAGCACTACCTCCAGAGAGGGGCGGTGAAGCG 352
 QY 114 AspProLysLeuSerLeuMetProTrpPheHisGlyLysIleSerGlyGlnGlnAlaVal 133
 353 GTACCAAACTACAGCCCTCATGCTGTGCTCCAGGCAAGATCAACAGGGAGAGCGCTGAG 412
 QY 134 GlnGlnLeuGlnProProGlnAspGlyLeuPheLeuValArgGlnSerAlaArgHisPro 153
 413 CGGCTTCTTACCCCGGAGACAGCCCTGCTGCTGGCGGAGACACCACTACCC 472
 QY 154 GlyAspTyrValLeuCysValSerPheGlyArgAspValIleHisTyrArgValLeuHis 173
 473 GGAGACTACACGCTGTGCTGAGCTCGAGCGGAGGAGGTGAGCACTACCGCATATGTAC 532
 QY 174 ArgAspGlyHisLeuThrIleAspGlyAlaValPhePheCysAsnLeuMetAspMetVal 193
 533 CATGCCAGCAAGCTCACATCGAGAGAGAGGTGATCTTGAGAACTCATGACAGCTGTG 592
 QY 194 GlnHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213
 593 GACCACTACACCTCAGCGAGAGATGTGACTGTACGGCCCTCATTAACCAAGGTATG 652
 QY 214 HisGlyThrLysSerAlaGlnGlnGlnLeuAlaArgAlaGlyTyrPheLeuAsnLeuGln 233
 653 GAGGGCACATGGGCGCCAGATGATTTCTACCGAGCGGCTGGCCCTGAACATGATAG 712
 QY 234 HisLeuThrLeuGlnAlaGlnIleGlyGlnGlyIlePheGlyAlaValLeuGlnGlyGln 253
 713 GACTGTGAAGCTGCGCAGACATCGGAGAGGGAGGTGCGAGACGTGATGCTGGCGAT 772
 QY 254 TyrLeuGlnGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
 773 TACCGAGGAAACAAATGCGCGTCAATGATTAAGCAACGCGCATCGCCAGCGCTTC 832
 QY 274 LeuAspGlnThrAlaValMetThrLysMetGlnHisGlnAsnLeuValArgLeuLeuGly 293
 833 CTGGCTAAGCCTCAGTCATGACGCAACTCGGCATGCAACCTGGTCAGCTCTGCGC 892
 QY 294 ValIleLeuHisGln-----GlyLeuTyrIleValMetGlnHisValSerLysGlyAsn 311
 893 GTGATCGTGAGAGAGAGGGCGGCTCTACATCGTCACTGATCAATGCCCAAGGGAGAGC 952
 QY 312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuLeuGln 331
 953 CTGTGGACTACCTGCGCTAGGGGTGCTGAGTGTGGCGGAGACTGCTCTCTCAAG 1012
 QY 332 PheSerLeuHisValAlaGlnGlyMetGlnTyrLeuGlnLysSerLysLeuValHisArg 351
 1013 TTCTCGCTAATGTCTCGAGAGCCATGAAATCTGAGAGCAACAATTTGCTGATCGA 1072
 QY 352 AspLeuAlaAlaArgAsnIleLeuValSerGlnAspLeuValAlaLysValSerAspPhe 371
 1073 GACTGGCTGCCCGCATGTGCTGTCTGAGAGCAACAGTGGCGCAAGTCAGGATTTT 1132
 QY 372 GlyLeuAlaLysAlaGlnArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391
 1133 GGCTCTCCACCAAGAGGCGTCCAGACCCAGGACCGGCGCAAGCTCCAGTCAAGTGGACA 1192
 QY 392 AlaProGlnAlaLeuLysHisGlyLysPheThrSerLysSerAspValTyrPsePheGly 411
 1193 GCCCTTAGGGCCCTGAGAGAGAAAGAAATTTCTCCACTAAGTCTGACGTGTGAGATTTCGA 1252

QY 412 ValLeuLeuTrpGlnValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLys 431
 1253 ATCTTCTTCGGGAAATCTACTCTTGTGGGAGATGCTTATCCAAATTCCTCTGAG 1312
 QY 432 GluValSerGlnAlaValGlnLysGlyTyrArgMetGlnProProGlnGlyCysProGly 451
 1313 GACTCGTCCCTCGCGGTGAGAGAGGCTACAGATGATGATGCCCGGAGCGCTGCCGCC 1372
 QY 452 ProValHisValLeuMetSerCysTrpGlnAlaGlnProAlaArgArgProPhe 471
 1373 GCAGTCTATGAGATCAAGAAACTGCTGGACCTGAGACCGCCATGCGGCTCTCTTC 1432
 QY 472 ArgLysLeuAlaGlnLysLeu 478
 1433 CTACAGCTCCGAGAGACGCTT 1453
 Db
 RESULT 9
 AAH28358
 ID AAH28358 standard; cDNA: 2187 BP.
 AC AAH28358;
 DT 05-SEP-2001 (first entry)
 DE
 XX Nucleotide sequence of human tyrosine kinase protein Src.
 KW Vascular permeability; tyrosine kinase protein; Src; Yes; stroke;
 KW myocardial infarction; restenosis; trauma; blood vessel; atherosclerosis;
 KW diabetic retinopathy; inflammatory disease; infection; arthritis;
 KW adult respiratory distress syndrome; ARDS; rheumatoid arthritis;
 KW diabetic retinopathy; psoriasis; neovascular glaucoma;
 KW capillary proliferation; osteoporosis; cancer; ss.
 KW
 XX Homo sapiens.
 OS
 XX
 FH Key
 FT CDS
 FT 134..1486
 FT /tag= a
 FT /product= "Src"
 XX
 XX W0200145751-A1.
 PD 28-JUN-2001.
 PD
 PF 22-DEC-2000; 2000WO-US35396.
 XX
 XX 22-DEC-1999; 99US-0470881.
 PR 29-MAR-2000; 2000US-0538248.
 XX
 XX (SCRI) SCRIPPS RES INST.
 PA
 PI Cheresh DA, Elliceiri B, Paul R;
 DR WPI: 2001-417982/44.
 DR P-PSDB: AAB84662.
 XX
 PT Modulating vascular permeability in tissues, including inflamed tissue,
 PT tissues associated with stroke, myocardial infarction, by contacting
 PT the tissue with tyrosine kinase protein Src, Yes or their modified
 PT forms -
 XX
 XX Disclosure; Fig 3; 133pp; English.
 PS
 XX The specification describes a method for modulating vascular
 CC permeability in a tissue suffering from a disease condition. The method
 CC comprises contacting the tissue with a pharmaceutical composition
 CC comprising tyrosine kinase protein Src, Yes or their mixtures or
 CC nucleic acid expressing them. The method is useful for modulating
 CC vascular permeability in tissues, including inflamed tissue, tissues
 CC associated with stroke, myocardial infarction or other blockage of
 CC normal flow, tissues undergoing restenosis, psoriatic, retinal tissue
 CC and similar tissues. Pathologies which may be treated include
 CC trauma to blood vessels, and other systemic pathological events such as

DR WPI: 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression
 of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -

XX Claim 1: SEQ ID NO 1173; 114bp; English.

CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs). Identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating Gs; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease, also bacterial infection, viral infection,
 CC parasitic infection, protozoal infection, fungal infection and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other:

Alignment Scores:

Pred. No.:	Length:	2187
Score: 1.16e-87	Matches: 238	
Percent Similarity: 71.59%	Conservative: 82	
Best Local Similarity: 53.24%	Mismatches: 118	
Query Match: 46.71%	Indels: 9	
	Gaps: 3	

US-09-977-260-2 (1-507) x ABK84602 (1-2187)

QY 40 AAlarqmetProthraTg-----ArgTPrAlaProGlyThGInCys 53

DB 116 GCCAGAGCTCTGGAGAGATGTCAGCAATACAGCCGCCCTGTCATCCGGTACGAATGT 175

QY 54 IlleThrLysGcSgLuHstHrArgProLysProGlyGluLeuAlaPheArgLysGlyAsp 73

DB 176 ATTCGCAAGTACCACTTCACAGGACATCCGACAGGACCTGCTCTGCAAGAGAGAC 235

QY 74 ValValThrIleLeuGluAlaCysGluAsnLysSerTrpTyrArgValLysHisHstHr 93

DB 236 GTGCTACACATTTGGCGCGTACACAGAGACCCCAACTGGTCAAAAGCAAAACAGATG 295

QY 94 SerLysGlnGluGlyLeuLeuAlaAlaGlyAlaLeuArgGlyLysGlyAlaLeuSerAla 113

DB 296 ---GGCGGTGAGGAGATCATCCAGCCAACTACGTCAGAGAGCGGAGGCGGTGAAGCG 352

QY 114 AsProLysLeuSerLeuMetProTrpPheHisGlyLysIleSerGlyGlnGluAlaVal 133

DB 353 GGTACCAAACTACAGCTCATCTGTTGTTCCACGCGCAAGATACACAGGAGACAGCTTAC 412

QY 134 GlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGlnSerAlaArgHisPro 153

DB 413 CGGCTCTGTACACCGCGGAGACAGAGCTGTCTGTGGTGGGAGAGACCAACTACCC 472

QY 154 GlyAspTyrValIleCysValSerPheGlyArgAspValIleHisTyrArgValLeuHis 173

DB 473 GGACACTACACAGCTGTGGTGTAGTGCACGCGCAAGGAGAGAGCTACCGCATCATGAC 532

QY 174 ArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVal 193

DB 533 CATGCCACAGAGCTACAGATGACAGAGAGGTGTACTTGGAAACCTCATCAGCTGCTG 592

QY 194 GluHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213

DB 593 GAGCAGTACACCTCAGACAGCAGATGAGACTGTACGCGCTCATTAACCAAGGTATG 652

QY 214 HisGlyThrLysSerAlaGlnGluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGln 233

DB 653 GAGGACACAGTGGCGGCCAGAGATGAGTTTACCGCACGCGCTGGCCCTCAACATGAG 712

QY 234 HisLeuThrLeuGlyAlaGlnIleGlyGlnGluPheGlyAlaValIleGlnGlyGlu 253

DB 713 GAGCTGAAGCTCTGTCAGACACCATCGGAGAGGGAGTGGAGACGTGATCTGGGCAT 772

QY 254 TyrLeuGlnGlyLysValAlaValAlaLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273

DB 773 TACGAGAGAACCAAAAGTGGCGGTCAAGTGAAGAACAGGACGACGCGCCAGGCTTC 832

QY 274 LeuAspGluThrAlaValMetThrLysMetClnHisGluAsnLeuValArgLeuGly 293

DB 833 CTGCTGAAGCTCATGATCATGACCACTGGGATGACCACTGGTCCACTCTGCGC 892

QY 294 ValIleLeuHisGln-----GlyLeuThrIleValMetGlnHisValSerLysGlyLys 311

DB 893 GTGATCTGGAGAGAGAGAGGCGGCTTACATCTGCTACTGATGATCATGCGCAAGGGAGC 952

QY 312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331

DB 953 CTGGGAGCTACCTCGGGGTATGAGGGGTGTCAGTGGTGGCGGAGACTGTCTCTCAG 1012

QY 332 PheSerLeuHisValAlaGlnGlyMetGlyTyrIleLeuGlnSerLysLeuValHisArg 351

DB 1013 TTCTCGCTAGATGTCCTGCGAGGCAATGAACTCGGAGGCAACAAATTTGTCATCGA 1072

QY 352 AsPheAlaAlaArgAsnIleLeuValSerGlnAspLeuValAlaLysValSerAspPhe 371

DB 1073 GACCTGCTGCCGCCCAATGTCTGTGTCTAGGACCAAGTGGCCCAAGTACGACATTT 1132

QY 372 GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391

DB 1133 GGCTCCACCAAGAGAGCGCTCCAGACCCAGACAGGCGCAAGCTGACGTCAGTGAAG 1192

QY 392 AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSerPheGly 411

DB 1193 GCCCTGAGGCCCTGAGAGAGAAATTCACATGATGTGACGTGTGAAGTTTCGA 1252

QY 412 ValLeuLeuTrpGluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLys 431

DB 1253 ATCTCTTCTGGAAATCTACTCTTTGGGAGAGTGCCTTATCCAAAGATTCCTGTAG 1312

QY 432 GluValSerGluAlaValGluLysGlyTyrArgMetGluProProGlnGlyCysProGly 451

DB 1313 GAGCTCGHCCCTCGGAGAGAGAGGCTACAAAGATGATGCCCGGAGGCTGCCGCC 1372

QY 452 ProValHisValIleMetSerCysTrpGluAlaGluProAlaArgArgProProPhe 471

DB 1373 GCACTTATGAAGATGAAAGAACTGCTGGACCTGGAGCGCCCATGCGCCTCTCTTC 1432

QY 472 ArgLysLeuAlaLulysLeu 478
 DB 1433 CTACAGCTCCGAGAGCAGCTT 1453
 RESULT 11
 ABL62918
 ID ABL62918 standard; DNA: 2187 BP.
 AC ABL62918:
 DT 15-MAY-2002 (first entry)
 DE Breast cancer related gene sequence SEQ ID NO:1255.
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 OS Homo sapiens.
 PN WO200194629-A2.
 PD 13-DEC-2001.
 PE 30-MAY-2001; 2001WO-US10838.
 XX 05-JUN-2000; 2000US-209473P.
 XX 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.
 PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 XX

PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 DR
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 1255; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 SQ Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.16e-87 Length: 2187
 Score: 1247.50 Matches: 238
 Percent Similarity: 71.59% Conservative: 82
 Best Local Similarity: 53.24% Mismatches: 118
 Query Match: 46.71% Indels: 9
 DB: 24 Gaps: 3
 US-09-977-260-2 (1-507) x ABL62918 (1-2187)
 QY 40 AlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
 DB 116 GCCAGAGCTTCAGAGATGTAGCAATACAGCCGCCCTGGCATCCGTTACAGATGT 175
 QY 54 IleThrLysCysGlnHisThrArgProLysProGlyGluLeuAlaHeArgLysGlyAsp 73
 DB 176 ATGCCAAGTACCAATCTCCAGCGACTGCCAGCAGCAGACTGCTCTGCACAAAGAGAC 235
 QY 74 ValValThrIleLeuAlaCysGluAsnLysSerTrpTrpArgValLysHisThr 93
 DB 236 GTGCTCACCATTTGTGGCCGTACACCAAGACCCCACTGTGTACAAACCAAAAGGTG 295
 QY 94 SerGlyGlnGluLeuLeuAlaGlyAlaLeuArgGluArgGluAlaLeuSerAla 113
 DB 296 ---GGCCGTGAGGATCATCCAGCCACACTACGTCCAGAGAGGAGGCGCTGAAGCG 352
 QY 114 AspProLysLeuSerLeuMetProTrpPheHisGlyLysIleSerGlyGlnGluAlaVal 133
 DB 353 GGTACCAAACTCAGCTCATGCTGTGTCTCTGTCTGTGGGAGACACCAACTCCCC 412
 QY 134 GlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgLysSerAlaGlnHisPro 153
 DB 413 CGGCTTCTGTACCGCGGAGACAGACGCTTCTGTCTGTCTGTGGGAGACACCAACTCCCC 472
 QY 154 GlyAspTrpValLeuCysValSerPheGlyArgAspValIleHisTrpArgValLeuHis 173
 DB 473 GGAGACTACAGCTGTGCTGAGCTCGACCGCAAGGTGAGAGCATACCGCATCATGTAC 532
 QY 174 ArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspVal 193
 DB 533 CATGCCAGCAGAGTACATCGACGAGGAGGTGTACTTTGAGAACTCATTCAGCTGTG 592
 QY 194 GlnHisTrpSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213

```

Db      593 GAGCACTACACCTGACGACGATGACTGTCTACGCGCTCATTAACCAAGTCATG 652
Qy      214 HisGlyThrIysSerAlaGluGluLeuAlaArgAlaGlyTTPLeuLeuLeuGln 233
Db      653 GAGGCGACAGTGGCGCCAGATGATCTTACCGCAGCGCTGGCCCTGACATGAG 712
Qy      234 HisLeuThrLeuGlyAlaGlnIleGlyGluGlyLeuPheGlyAlaValLeuGlnGlyGlu 253
Db      713 GAGCTGAAGCTGCTGCAGACCATTCGGAGAGGGAGTTCGGAGACGTCGATCTGGCGAT 772
Qy      254 TyrLeuGlyGlnValAlaValAlaLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
Db      773 TACCGAGGAAACAAGTGCCTGCAATGCTTAAGAAGCAGCGACCTGCCAGCGCTTC 832
Qy      274 LeuAspGluThrAlaValMetThrLysMetGlnHisGluAsnLeuValArgLeuLeuGly 293
Db      833 CTGGCTAAGCTAGTCATGACCAACTGGCGCATGCAACCTGGTGCACCTCTCTGGGC 892
Qy      294 ValIleLeuHisGln-----GlyLeuTyrIleValMetGluHisValSerLysGlyAsn 311
Db      893 GTGATCTCGAGAGAGAGGCGGCTCTACTCTCTCATCTGCTCATGTCATGCGCAAGGGAGGC 952
Qy      312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuLeuGln 331
Db      953 CTCTGGACTACCTGCGGTCTAGGGGTGCGTCACTGCTGGCGGAGACTCTCTCTCAAG 1012
Qy      332 PheSerLeuHisValAlaGluGlyMetGluTyrLeuGluLysSerLysLeuValHisArg 351
Db      1013 TTCCTCGTACATGTCTCGAGGCGCATGTAATCCTGAGGCAACATTTGCTCATCGA 1072
Qy      352 AspleuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371
Db      1073 GACTGTGCTGCGCCCATGTCTGTCTAGAGCAACAGTGGCCAGCTCAGGACTTT 1132
Qy      372 GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391
Db      1133 GCTCTCACCAAGGAGGCGTCCAGCACCCAGCAGCGGCAACGTCGACATCAAGTGGACA 1192
Qy      392 AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSerPheGly 411
Db      1193 GCCCTTAGGCGCTGAGAGAGAGAAATCTCCACTAAGTGTGAGGTGAGTTGGGA 1252
Qy      412 ValLeuLeuTrpGluValPheSerTyrGlyArgAlaProLysProLysMetSerLeuLys 431
Db      1253 ATCTTCTCTGGGAAATCTACTCTTTGGGGAGTGCCTTATCCAGATTCCTCTGAG 1312
Qy      432 GluValSerGluAlaValGluLysGlyTyrArgMetGluProProGluGlyCysProGly 451
Db      1313 GACCTCTCCCTCCGCGGAGAAAGGCTACAAAGATGATGCCCGGAGCGCTGCCGCC 1372
Qy      452 ProValHisValLeuMetSerSerCysTrpGluAlaGluProAlaArgProProPhe 471
Db      1373 GCACCTTATGAAGTCAATGAACTGCTGCGACCTGAGCGCCGATGCGGCTCTCTTC 1432
Qy      472 ArgLysLeuAlaGluLysLeu 478
Db      1433 CTACAGCTCCGAGAGGACGCTT 1453

```

RESULT 12

AA246489 standard; DNA; 2420 BP.

AA246489;

13-MAR-2000 (first entry)

PKA substrate, Csk-family protein encoding DNA.

Protein kinase A; PKA; PKA signaling pathway; phosphorylation; cancer;
 kinase substrate; immunosuppressive disorder; proliferative disease;
 HIV infection; AIDS; immunodeficiency; autoimmune disease; Chk; Lsk; Hyl;
 systemic lupus erythematosus; Csk-family; Matk; Ctk; Bnk; Ntk; ss.

```

XX      XX Homo sapiens.
XX      OS
XX      FH Key Location/Qualifiers
XX      FT CDS 413..1765
XX      FT /tag= a
XX      PN MO9962315-A2.
XX      PD 02-DEC-1999.
XX      PE 27-MAY-1999; 99WO-GB01680.
XX      PR 27-MAY-1998; 98NO-0002419.
XX      PR 30-DEC-1998; 98US-0114240.
XX      PA (LAUR-) LAURAS AS.
XX      PA (JONE-) JONES E L.
XX      PI Hansson V, Levy FO, Mustelin T, Skalhogg BS, Sundvold V, Tasken K;
PI      Vang T, Altman A, Munshi A;
DR      WPI: 2000-086801/07.
DR      P-PSDB: AAY49418.
PT      Altering the activity of protein kinase signaling pathways, used for
PT      treating immunosuppressive disorders, e.g. AIDS, proliferative
PT      disorders, e.g. cancers or autoimmune diseases
PS      Claim 11: Page 87-89; 11pp; English.
XX      XX
XX      XX The invention provides a novel method of altering the activity of the
XX      XX protein kinase A (PKA) signaling pathway in a cell that comprises
XX      XX altering the extent of phosphorylation of one or more PKA substrates, or
XX      XX kinase substrates downstream in the PKA signaling pathway. Pharmaceutical
XX      XX compositions containing a nucleic acid molecule that encodes a PKA
XX      XX substrate, or fragment, precursor or functionally equivalent variant,
XX      XX where the sequence is modified to alter its susceptibility to
XX      XX phosphorylation by PKA can be used for treating a disorder exhibiting
XX      XX abnormal PKA signaling activity, immunosuppressive disorders or
XX      XX proliferative diseases. They can be used for treating e.g. HIV
XX      XX infection, AIDS, common variable immunodeficiency or cancers. Conditions
XX      XX in which upregulation of the PKA pathway is required, such as autoimmune
XX      XX disease, e.g. systemic lupus erythematosus, may also be treated. The
XX      XX present sequence represents a DNA sequence encoding a PKA substrate,
XX      XX wherein the substrate is in the Csk-family, preferably Csk, Chk, Lsk,
XX      XX Hyl, Matk, Ctk, Bnk or Ntk.
XX      XX
XX      XX Sequence 2420 BP; 471 A; 771 C; 703 G; 475 T; 0 other;
SQ
XX      XX
XX      XX Alignment Scores:
XX      XX Pred. No.: 1.32e-87 Length: 2420
XX      XX Score: 1247.50 Matches: 238
XX      XX Percent Similarity: 71.59% Conservative: 82
XX      XX Best Local Similarity: 53.24% Mismatches: 118
XX      XX Query Match: 46.71% Indels: 9
XX      XX DB: 21 Gaps: 3
XX      XX
XX      XX US-09-977-260-2 (1-507) x AA246489. (1-2420)
XX      XX
Qy      40 AlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
Db      395 GCCAGAGCTCTGAGAGATGTCAGCAATACAGGCCCGCTGCGATCGGATGATGT 454
Qy      54 IleThrLysCysGluHisThrArgProLysProGlyLeuLeuAlaPheArgLysGlyAsp 73
Db      455 ATTGCCAAGTACCACTTCCAGCGCATGCCGAGCAGACCTGCTTCTGCAAGGAGAC 514
Qy      74 ValValThrIleLeuGluAlaCysGluAsnLysSerTyrArgArgValLysHisThr 93
Db      515 GTGCTCACCATTTGGCGCGTCCACAGAGACCCCACTGATGTCACCAAGCAAGAGGTG 574
Qy      94 SerGlnGlnGlyLeuLeuAlaGlyAlaLeuArgGluArgGluAlaLeuSerAla 113

```

```

Db 575 ---GGCGGTGAGGCGATCATCCAGCCACTACGTCCAGAGAGCGGAGGCGGTGAAGCG 631
Qy 114 AspProLysLeuSerMetProTrrPheHisGlyLysIleSerGlyInGluAlaVal 133
Db 632 GGTACCAACTCAGACCTCATGCTGTGGTCCAGGCAAGATCAGAGGAGAGCGGTGAG 691
Qy 134 GlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGluSerAlaArgHisPro 153
Db 692 CGGTTTGTGACCGCGGAGACAGCGCTGTCTGTGGGAGACACCAACTATACCCC 751
Qy 154 GlyAspTyrValLeuGlyValSerPheGlyArgAspValIleHisTyrArgValLeuHis 173
Db 752 GAGACTACACAGCTGTCTGAGCTGAGCTGAGCGGCAAGGTGAGAGCTACCCATCATATAC 811
Qy 174 ArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVal 193
Db 812 CATGCCAGCAAGCTCAGACATCGAGAGAGGTGTACTTGAACCATGACATGACAGTGTG 871
Qy 194 GlnHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213
Db 872 GACACATACACCTCAGACGAGATGAGCTGTACGCGCTCATTAACCAAGGTCATG 931
Qy 214 HisGlyThrLysSerAlaGlnGluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGln 233
Db 932 GAGGCGACAGTGGCGGCGGAGATGATTTACCGGAGCGGCTGGCCCTGACATGAG 991
Qy 234 HisLeuThrLeuGlyAlaGlnIleGlyGluGlyIlePheGlyAlaValLeuGlnGly 253
Db 992 GAGCTGAAAGCTGCGGAGACATCGGAGAGGAGGTGGAGAGCGTGAATGCGGCGAT 1051
Qy 254 TyrLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
Db 1052 TACCGAGGAAACAAGTGGCGCTCAAGTGTAAAGAACACCGCCATGCCAGCGCTTC 1111
Qy 274 LeuAspGluThrAlaValMetThrLysMetGlnHisGluAsnLeuValArgLeuLeuGly 293
Db 1112 CTGGCTGAAAGCTCAGTACATGACGCACTCGGATAGCAACCTGGGAGCGCTCGGCG 1171
Qy 294 ValIleLeuHisGln-----GlyLeuTyrIleValMetGlnHisValSerLysGlyAsn 311
Db 1172 GTGATGCTGAGAGAGAGGCGGCTGTACATGCTGCTACATGACATGCGCAAGGAGCG 1231
Qy 312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
Db 1232 CTGTGTGACATCACTGCGGTCTAGCGGTGCTGAGTGTGCGGAGACTGTCTCTCAAG 1291
Qy 332 PheSerLeuHisValAlaGlnGlyMetGluTyrLeuGluSerLysLysLeuValHisArg 351
Db 1292 TTCTCGCTAGATGTCTGCGAGGCAATGGAATACCTGGAGGCAACAATTTCTGTCATCGA 1351
Qy 352 AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371
Db 1352 GACCTGGCTCCCGCATGTGTGTGTGTGAGAGACACTGGCCAAAGTCACAGCACTTT 1411
Qy 372 GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerArgLeuProValLysTrrPhe 391
Db 1412 GGCTCAACCAAGAGCGCTCACACCAGCGGCAAGCTGCGACATGCAAGTGGACA 1471
Qy 392 AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrrPheSerPheGly 411
Db 1472 GCGCCGAGGCGCTGAGAGAGAGAAATTTCTCACATAGCTCATAGCTGTGAGATTTCGGA 1531
Qy 412 ValLeuLeuTrrPgluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLys 431
Db 1532 ATCTCTCTCGGAAATCTACTCTTGGCGAGTGCCTTATCCAGAAATTCCTCCGTAAG 1591
Qy 432 GluValSerGluAlaValGluLysGlyTyrArgMetGluProProGluGlyCysProGly 451
Db 1592 GAGCTGCTCCCTCGGTGAGAGAGGCTCAACATGATGATGCCCCACAGCGCTCCGCC 1651
Qy 452 ProValHisValLeuMetSerCysTrrPgluAlaGluProAlaArgArgProPhe 471

```

```

Db 1652 GCACTCATATAGATGAAGAACTGCTGCGACCTGAGACCGCCGATCGGCCCTCTTC 1711
Qy 472 ArgLysLeuAlaGluLysLeu 478
Db 1712 CTACAGCTCCGAGAGCAGCTT 1732

RESULT 13
ACC50120
ID ACC50120 standard; cDNA; 2420 BP.
XX
AC ACC50120;
XX
DT 12-JUN-2003 (first entry)
XX
DE Breast cancer associated cDNA sequence SEQ ID NO:87.
XX
KW Human; breast cancer; cytostatic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200304989-A2.
XX
PD 16-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US19669.
XX
PR 21-JUN-2001; 2001US-299887P.
PR 27-JUN-2001; 2001US-301572P.
PR 18-JUL-2001; 2001US-306501P.
PR 25-SEP-2001; 2001US-325002P.
PR 05-MAR-2002; 2002US-362585P.
PR 14-MAY-2002; 2002US-380391P.
XX
PA (MILL-) MILENIUM PHARM INC.
XX
PI Lillie J, Gannavarapu M, Ghatt K, Hoersch S, Kamatkar S, Mertens M;
PI Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE, Bast RC;
PI Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
DR WPI; 2003-210381/20.
XX
P P-PSDB; ABR47428.
XX
PT Breast cancer diagnosis or treatment by comparing the level of
PT expression of a marker in a patient sample with that in the control
PT non-breast cancer sample
XX
PS Claim 1; SEQ ID 87; 128pp: English.
XX
CC The present invention describes a method for assessing whether a patient
CC is afflicted with breast cancer. The method comprises comparing the level
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC ABR47386 to ABR47632) in a patient sample and the normal level of
CC expression of the marker in a control non-breast cancer sample, where a
CC significant increase in the level of expression of the marker in the
CC patient sample and the normal level is an indication that the patient is
CC afflicted with breast cancer. The breast cancer associated sequences
CC from the present invention have cytostatic activities and can be used in
CC gene therapy. The method is useful for diagnosing and treating breast
CC cancer.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 2420 BP; 471 A; 771 C; 703 G; 475 T; 0 other;
XX

Alignment Scores:
Pred. No.: 1,32e-87 Length: 2420
Score: 1247.50 Matches: 238
Percent Similarity: 71.59% Conservative: 82
Best Local Similarity: 53.24% Mismatches: 118
Query Match: 46-71# Indels: 9
DB: Gaps: 3

```

US-09-977-260-2 (1-507) x ACC50120 (1-2420)

QY 40 AlaArgMetProThrArg-----ArgTrpAlaProGluYThrcIncs 53
 DB 395 GCCACAGCTCTGAGAAAGATGTCACCAATACAGGCCCTGGCCATCCGGTACAGAAAGT 454
 QY 54 IleThrLysCysGlnHisThrArgProLysProGluLeuAlaPheArgLysLysPhe 73
 DB 455 ATTGCCAAGTACAACTCCACGGGCACTCCGAGCAGCACTGCCCTTGTGCAAGAGAAC 514
 QY 74 ValValThrIleLeuGluAlaCysGlnLysSerTrpPylArgValLysHisThr 93
 DB 515 GTGCTACCACTATGTTGGCGTACCAAGGACCCCACTGTACAAAGCAAAACCAAGGTG 574
 QY 94 SerGlyGlnGluLeuLeuAlaGluAlaLeuArgGluArgGluAlaLeuSerAla 113
 DB 575 ---GGCGGTGAGGGGATCATCCACCACTACCTCCGAGGAGGAGGCGGTGAGGCG 631
 QY 114 AspProLysLeuSerLeuMetProTrpPheHisGlyLysIleSerGlyGlnGluAlaVal 133
 DB 632 GGTACCAAACTCAGCTCATGCTGTTCCACGGCAGATCACACGGGAGCAGGCTGAG 691
 QY 134 GlnGlnLeuGlnProGluLysPylLeuPheLeuValArgGluSerAlaArgHisPro 153
 DB 692 CGGCTTCTGTACCCGCGGAGACAGGCTGTCTGCGGAGAGCAGCACTACATCC 751
 QY 154 GlyAspPylValLeuCysValSerPheGlyArgAspValIleHisThrArgValLeuHis 173
 DB 752 GGAAGCTACACGCTGTGCTGAGCTGCGAGCAGCAAGGAGGAGCACTACCATCATGAC 811
 QY 174 ArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVal 193
 DB 812 CATCCACAGCAAGCTCAGATGAGAGAGGAGGTGACTTTGGAACCTCATCATCAGCTG 871
 QY 194 GluHisTrpSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213
 DB 872 GAGCAGTACACCTCAGACGCGAGATGACTCTGTACGCGCCCTATTAAACCAAGGTCATG 931
 QY 214 HisGlyThrLysSerAlaGluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGln 233
 DB 932 GAGGCGACAGTGGCGCCAGATGATGTTCACGAGCGGCTGGCCCTGACATGAG 991
 QY 234 HisLeuThrLeuGlyAlaGlnIleGlyGlnGlyLysLysGlyAlaValLeuGlnGlyGlu 253
 DB 992 GAGCTGAGCTGCTGCGACACCATCGGAGAGGAGGAGTTCGAGACGTCGATGCTGGCGAT 1051
 QY 254 TyrLeuGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
 DB 1052 TACCGAGGAAACAAAGTCGCGTCATGATTAAGAACAGCAGCCACTGCCAGGCTTC 1111
 QY 274 LeuAspGluThrAlaValMetThrLysMetGlnHisGlnLysLeuValArgLeuGly 293
 DB 1112 CTGGCTCAGAGCTCAGTCATGACCACTGCGGATGACCAAGCTGGAGCTCTCGGCG 1171
 QY 294 ValIleLeuHisGln-----GlyLeuThrIleValMetGlnHisValSerLysGlyAsn 311
 DB 1172 GTGATCTCTGGAGGAGGAGGCGGCTCTACATGCTCACTAGTACATGAGCCAAAGGAGC 1231
 QY 312 LeuValAsnPheLeuArgTrpArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
 DB 1232 CTGTGGACTACCTGCGGCTAGGCGTCAAGTGTGGCGGAGAGCTGTCTCTCAAG 1291
 QY 332 PheSerLeuHisValAlaGluGlyMetGlnTyrLeuGlnLysSerLysLeuValHisArg 351
 DB 1292 TTCCTGCTGATGCTGCGAGGCGCATGAAATACCTGAGCAACAAATTCGTCATCGA 1351
 QY 352 AspLeuAlaAlaArgAsnIleLeuValSerGlyAspLeuValAlaLysValSerAspPhe 371
 DB 1352 GACTGGCTGCCCGCATGTGCTGCTGAGGAGCAACAGTGGCCAAAGGTCAGGACTTT 1411
 QY 372 GlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391
 DB 1412 GGTCTCACCAGAGGCGTCCAGCACCAGCAGCAGGCAAGCTGCCAGTCAAGTGGACA 1471

QY 392 AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSerPheGly 411
 DB 1472 GCCCTGAGGCCCTGAGAGAGAGAGAAATTCACACTAAGTGTGACGTGGAGTTTCGA 1531
 QY 412 ValLeuLeuTrpGluValPheSerTrpGlyArgAlaProTrpProLysMetSerLeuLys 431
 DB 1532 ATCCTCTCTGGAAATATCTACTCTTGGCGGAGTGCCTTATCCAAAGATTCCTCGAAG 1591
 QY 432 GluValSerGluAlaValGluLysGlyTyrArgMetGluProProGluGlyCysProGly 451
 DB 1592 GACGTGCTCCCTCGGTGAGAGAGGCTACAAAGATGATGCCCGGAGGCTGCCGCC 1651
 QY 452 ProValHisValLeuMetSerSerCysTrpGluAlaGluProAlaArgArgProPhe 471
 DB 1652 GCAGTCTATGAAAGTCATGAAAGACTGTGGCAGCCTGGAGCCGCCATGCGCCCTCCTTC 1711
 QY 472 ArgLysLeuAlaGluLysLeu 478
 DB 1712 CTACAGCTCCGAGAGCAGCTT 1732
 DB
 RESULT 14
 AA193804
 ID AA193804 standard; cDNA; 2466 BP.
 XX
 AC AA193804;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 13864.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 OS Homo sapiens.
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 DR WPI: 2001-514838/56.
 DR P-PDB: AA013873.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 XX disorders -
 PS Claim 1; SEQ ID NO 13864; 1399pp + Sequence Listing; English.
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2466 BP; 483 A; 785 C; 715 G; 483 T; 0 other;

Alignment Scores:

Pred. No.:	1.35e-87	Length:	2466
Score:	1247.50	Matches:	238
Percent Similarity:	71.59%	Conservative:	82
Best Local Similarity:	53.24%	Mismatches:	118
Query Match:	46.71%	Indels:	9
	22	Gaps:	3

US-09-977-260-2 (1-507) x AA193804 (1-2466)

```

QY 40 AAlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
   ||||| ||| ::| ||||| ||||| ||||| ||||| ||||| |||||
Db 395 GCCAGAGCTCTCGAGAAATGTCAGCAATACAGCGCGCTGGCAGTCCGGTACAGAAATGT 454
QY 54 TleThrLysCysGlnHisThrArgProLysProGlyGluLeuAlaPheArgLysGlyAsp 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 ATTCGCCAAGTACAACTTCCAGCGCACTGCCGAGCAGACCTGCCCTTCTGCAAAAGAGAGAC 514
QY 74 ValValThrIleLeuGlnAlaCysGluAsnLysSerTrpTrpArgValLysHisThr 93
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 515 GTGCTACCACTTGTGGCGCTCACACAGAGACCCCACTGTGACAAACCAAAACAAGTGG 574
QY 94 SerGlyGlnGluGlyLeuLeuAlaAlaGlyAlaLeuArgGluLysGluAlaLeuSerAla 113
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 575 ---GGCGGTAGAGGCGATCCACAGCACTACGTCCAGAGGGGAGGGGCGTGAAGGCG 631
QY 114 AspProLysLeuSerLeuMetProTrpPheHisGlyLysIleSerGlyGlnGluAlaVal 133
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 632 GGTACCAAACTACAGCTCATGCTTGTTCACGGCAATCATCACAGCGAGCGAGCTGAG 691
QY 134 GlnGlnLeuGlnProProGlnAspGlyLeuPheLeuValArgGluSerAlaArgHisPro 153
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 692 CGGCTTCTGTACCCCGCGGAGACAGGCGCTTCTGTCGGGAGAGCAACACTACCC 751
QY 154 GlyAspTrpValLeuGlyValSerPheGlyArgAspValIleHisTrpArgValLeuHis 173
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 752 GGAGACTACACCTGTGCTGAGCTGCAGCGCAAGGAGGAGACACTACCGCATCATGTCAG 811
QY 174 ArgAspGlyHisLeuThrIleAspGlyAlaValPhePheCysAsnLeuMetLysPheVal 193
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 812 CATGCCAGCAACTCTGACATCGACAGGAGGTGCTTGTGAGAACTCATGACGCTGAG 871
QY 194 GlnHisTrpSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLysArgLys 213
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 872 GAGCACTACACTCAGACGACAGATGAGACTGTGACGCGCTCATTAACCAAAAGGTCATG 931
QY 214 HisGlyThrLysSerAlaGluGlnGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGln 233
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 932 GAGGGCAGACGTGCGCCAGATGATGTCACCGAGCGCGCTGAGCGCTGAACATGAG 991
QY 234 HisLeuThrLeuGlyAlaGlnIleGlyGlnGlyGlnPheGlyAlaValLeuGlnGlyGlu 253
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 992 GAGCTGAAGCTCTCGACACCATCGGAGGAGGAGATTGGAGACCTGATGCTGGCGCAT 1051
QY 254 TyrLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1052 TACCGAGGAGAACAAAGTCCCGCTCAAGTCAATTAAGACGAGCCCTCCAGGCGCTTC 1111
QY 274 LeuAspGluThrAlaValMetThrLysMetGlnHisGluAsnLeuValArgLeuGlnGly 293
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1112 GTGGCTGAGCCTCAAGTCAAGCAACTGGCGGCATAGCACTGTGTCAGCTCCGGGC 1171
QY 294 ValIleLeuHisGln-----GlyLeuTyrIleValMetGlnHisValSerLysGlyAsn 311
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1172 GTGATCTGTGAGAGAGGCGCGGCTCTACATCGTCACAGAGTACTGTGCGCAAGGCGAC 1231
QY 312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1232 CTGTGTGACTACTGCGGTCTAGGGGTGCGTCACTGCTGGCGGAGACTGTCTCTCAAG 1291

```

```

QY 332 PheSerLeuHisValAlaGlnGlyMetGluTyrLeuGlnSerLysLysLeuValHisArg 351
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1292 TTCTGCTAGATGTGTGCGAGGCGCATGAGAACTGAGGAGCAAAATTCGTGATGCA 1351
QY 352 AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerPhe 371
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1352 GACCTGGCGCGCCGCAATGTGCTGTGAGGACAGACGTGCCAAAGCTCAGCGACTTT 1411
QY 372 GlyLeuAlaLysAlaGlnArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1412 GTCTCACCAAGAGGCGCTCCAGACCCAGACAGCGGCAAGCTGCCAATCAAGTGAACA 1471
QY 392 AlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSerPheGly 411
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1472 GCCCTGAGCGCCCTGAGAGAGAAATTCCTCACTAAGCTGACGTGTGAGATTTCGGA 1531
QY 412 ValLeuLeuTrpGluValPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLys 431
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1532 ATCCCTTCTGTGGAAATCTACTCTTGGCGAGTGCCTTATCCAAAGAAATTCCTCGAAG 1591
QY 432 GluValSerGluAlaValAlaGlnLysGlyTyrArgMetGluProProGlnGlyCysProGly 451
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1592 GACGTGCTCCCTCGGCTGAGAGAGGCTTACAGATGATGCCCGCAGCGCTGCCGCC 1651
QY 452 ProValHisValLeuMetSerSerCysTrpGluAlaGlnProAlaArgArgProPhe 471
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1652 GCAGTGTATGAGTGTATGAAGAACTGCTGCACTGAGCGCGCGCATGGCGCTTC 1711
QY 472 ArgLysLeuAlaGlnLysLeu 478
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1712 CTACAGCTCCGAGAGCAGCTT 1732

```

RESULT 15

ABS73326

ID ABS73326 standard; DNA; 2442 BP.

XX ABS73326;

AC ABS73326;

DT 04-DEC-2002 (first entry)

XX

DE DNA encoding human c-src isoform.

XX

KW Chromosome aberration; oncogenic fusion protein; cancer; oncogene;

KW proliferative disease; cellular protein isoform; heat shock protein 90;

KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;

KW T cell lymphoma; B cell lymphoma; chronic myeloid leukemia; CML;

KW acute myeloid leukemia; AML; chronic myelomonocytic leukemia; CMML;

KW acute lymphoblastic leukemia; ALL; APL; NHL; solid tumour;

KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;

KW rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.

XX Homo sapiens.

XX

PN WO200269900-A2.

XX

PD 12-SEP-2002.

XX

PF 01-MAR-2002; 2002MO-US06518.

XX

PR 01-MAR-2001; 2001US-272751P.

XX

PA (CONF-) CONFORMA THERAPEUTICS CORP.

XX

PI Fritz LC; Burrows FJ;

XX

DR WPI; 2002-698710/75.

XX

XX P-PSDB; ABG95117.

PT Treating genetically-defined disease associated with chromosomal

PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative

PT diseases, involves administering an inhibitor of heat shock protein 90

xx Disclosure; Page 313-314; 389pp; English.
 CC The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (I) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant
 CC protein or cellular protein isoforms (II) dependent on heat shock
 CC protein (HSP)-90, or selectively treating cells expressing (II)
 CC involving administering HSP90-inhibitor. The method is useful for
 CC treating genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, MLL, NHL and CML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This sequence encodes a human oncogenic protein.
 CC
 xx Sequence 2442 BP; 474 A; 779 C; 708 G; 481 T; 0 other;

Alignment Scores:

Pred. No.: 1,266-83 Length: 2442
 Score: 1196.50 Matches: 234
 Percent Similarity: 72.418 Conservative: 81
 Best Local Similarity: 53.798 Mismatches: 116
 Query Match: 44.80% Indels: 5
 DB: 24 Gaps: 2

US-09-977-260-2 (1-507) x ABS73326 (1-2442)

QY 47 TTPALAPROGLYTHRGNGLYSLHLEPHLYSCYSGIUNHISHTHARGPROLYSGIUN
 DB 460 TGGCCATCCGGTACAGATGTATTCGCAAGTACACTCCACGGGACGCGAGAGGAC
 QY 67 LEUALPHEATRGVGLYASPVAYALHTRILEUGIUALCYSGIUNSLYSETRTP
 DB 520 CTGCCCTCTGCAAGGAGAGCGTGCTCACTGTCGCGCCGACCAAGGACCCCAACTGG
 QY 87 TYRARGVALYSHHSHTHSERGLYNGIUNGLYLEUVALAALYALAEUARG 106
 DB 580 TACAAAGCCAAAGAGAGTG---GGCCGTAGGAGCATCATGCCACCAACTACGTCCAG
 QY 107 GIUARGIUALAEUSERIALASPRPROLYSEUSERLEUMETPRIO-TRYPHENISGLY 126
 DB 637 AAGCGGAGGGGGGTAGAGGGGTACCAAACTCACCTCATGCCGTGAGTTCCAGGCA
 QY 126 SILESERGLYNGIUALAVALGNGIUNGLINPROFROGLUASPRGLYLEUPHEUVA 146
 DB 697 GATCACAGCGGAGCGGTAGCGGCTGTCTTACCCGCGGAGACAGCGCTGTCTGT
 QY 146 IARGIUSERIALASPRPROGLYASPTVALYLEUCYSVALSERPHEGLYARGASPYA 166
 DB 757 GCGGAGAGGACCAACTACCCCGAGACTACACGCTGTGCTGAGCTGCGAGCGCAAGT
 QY 166 ILEHISTYUARGVALYLEUNHISARGASPRGLYHISLEUTHRIEASPRGLUALAVALPHER 186
 DB 817 GAGAGCATACCGCATCATGATCCATGCGACAGCACTACAGCATGAGAGAGGTACTT
 QY 186 ECYASALEUWETASPMETVALGUNGISTYUSETLYSASPRLYGIALALIECYSTHLY 206
 DB 877 TGAGAACCTCATGCGCTGTGAGACACTACACTCAGACGCGAGATGAGTCTGTACGCG
 QY 206 SLEUVALARGPROLYSARGVLYSHISGLYTHLYSSEIALAGIUNGLIUEUALAARGAL 226
 DB 937 CCTCATTTAAACCAAGGATGAGAGGACAGTGGGCGGAGATGAGTTTACCGAG 996
 QY 226 AGLYTRPHEUVALAEUNGLINHISLEUTHRIEUGIUALAGIUNGLIUGIUNH 246

DB 997 CGCGTGGGCGCTGAACATGAGAGAGCTGAAGCTGCTGCAGACCACTCCGAGAGGGAGT 1056
 QY 246 EGIYALVALLEUNGINGIUTRYLEUGIUNGLYVALALVALYLSANILIELYCY 266
 DB 1057 CGGAGACGTGATGCTGGGCGATTCACCGAGGAGCAAAATCCCGTCAAGTCAATTAA 1116
 QY 266 SASPVATHRIAGIUNHIALPHEUASPRGLIUTHRIALVALMETHRLYSEMETGLINHSGL 286
 DB 1117 CGAGGCCACTGCGCCGCTTCTGCGTGAAGGCTCACTCATAGCAGCACTCGCGCATAG 1176
 QY 286 UASPVATHRIAGIUNHIALPHEUASPRGLIUTHRIALVALMETHRLYSEMETGLINHSGL 304
 DB 1177 CAACCTGTGACAGCTCTCGGCGGTGATGTGAGAGAGAGGCGGCTCTACATCTGTAC 1236
 QY 304 TGLIUNHISVALSERLYSGIUNGLYASPVAYALASPRPHEUARGHTRARGVALAARGALAEUVA 324
 DB 1237 TGAGTACATGGCCCAAGGGAGCGCTGTGAGTACTGCGGTGAGGCTGAGTGTGCT 1296
 QY 324 IASNTHRIAGIUNHIALPHEUASPRGLIUTHRIALVALMETHRLYSEMETGLIUTRYLEUGI 344
 DB 1297 GCGCGGAGACGTCTCTCCATGATTTCTGCTGATGATGTCTGCGAGGCGCATGGAATCTGGA 1356
 QY 344 USERLYSLYSEUVALHISARGASPRLEUALAARGASNILEUVALSERGLIASP 364
 DB 1357 GGGCAACAATTTCTGTCATCGACATCGAGACTGCTGCCCGCAATGTGCTGTGAGAGCA 1416
 QY 364 UVALALALYVALSERASPRPHEGLYLEUALALYSLAGIUNGLYSGIUNLYSEASPRSE 384
 DB 1417 GGTGGCAAGGTGACGAGCACTTTGCTGTCCACCAAGGAGCGCTCCACACC-CAAGACACGGG 1475
 QY 384 FARGUEUPROVALYSTRPHTHRIAPROGLIUALAEUYSHISGLYSPHEHTRSELY 404
 DB 1476 CAAGCTGCCAGTCAAGTGGACAGCCCTGAGGCGCTGAGAGAGAGAAATTTCTCCACAT 1535
 QY 404 SSERASPVATHRISSERPHEGLYVALLEUETHRPGIUNVALPHESETRYGIALAARG 424
 DB 1536 GTCTGACGTGTGAGATTTCTGCAATCTTCTGTGGAAATCTCTCTTGGGCGAGTCC 1595
 QY 424 CTYTRPROLYSMETSERILEUVALYVALSERGLIUALAVALGIUNYSGIUTYARGMETGL 444
 DB 1596 TTATTCAGAAATTCCTGTGAGAGAGCTGTGCTGCGGTGGAGAGGCTGTACAAAGTGA 1655
 QY 444 UPROFROGLIUCYSPROGLYPROVALHISVALLEUMETSERCYSTRPGIUALAGI 464
 DB 1656 TGCCCCCAAGCGCTGCCCGCGGAGCTGTATGATCATGAAGAACTGTGCGACACTGGA 1715
 QY 464 UPROVALARGARGPROFROPHARGVLYSLEUALAGIUNLYSEU 478
 DB 1716 CGCCGCAWTGGCGGCTCTCTTCTTACAGCTCCGAGACAGAGCTT 1758

Search completed: August 1, 2003, 20:02:54
 Job time : 504 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2003, 18:16:35 ; Search time 6032 Seconds

(without alignments)
3438.525 Million cell updates/sec

Title: US-09-977-260-2
Perfect score: 2671
Sequence: 1 MAGRSLVSMRAFHCDSAE.....PASVSGDADGSTSPRSGEP 507

Scoring table: BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2688711 segs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO/spool/US09977260/runat_29072003_092741_14619/app.query.fasta.1.647
-DB=genEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09977260.ecgn.1.1.5265.ernat.29072003_092741_14619 -NCPU=6 -ICPU=3
-NO_MMAR -LARGEDUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2671	100.0	1942	6 AR053292	AR053292 Sequence
2	2671	100.0	1942	6 E10345	E10345 cDNA encodi
3	2671	100.0	1968	6 HSHYLTJK	X77278 H sapiens H
4	2671	100.0	1997	6 BD176710	BD176710 Method of
5	2671	100.0	2000	6 AR179662	AR179662 Sequence
6	2671	100.0	2000	6 AR274487	AR274487 Sequence
7	2671	100.0	2180	6 BC000114	BC000114 Homo sapi
8	2671	100.0	2180	6 BC003109	BC003109 Homo sapi
9	2662	99.7	1521	6 AR053291	AR053291 Sequence
10	2607.5	97.6	1744	6 S71669	S71669 1st-Leukocy
11	2591.5	97.0	1987	6 AR084686	AR084686 Sequence
12	2591.5	97.0	1987	6 AR204704	AR204704 Sequence
13	2591.5	97.0	1987	6 HSDMATRK	LI18974 Human tyros
14	2544	95.2	1804	9 HSM802388	AL13754 Homo sapi
15	2541.5	95.2	2073	6 AK055395	AK055395 Homo sapi
16	2444	91.5	1398	6 AR053290	AR053290 Sequence
17	2341	87.6	1734	10 MDSNTR	L27738 Mus musculu
18	2290.5	85.8	1818	10 MDSCTK	D45242 Mouse mRNA
19	2279.5	85.3	1931	10 RATABTK	L34542 Rattus norv
20	2277.5	85.3	1656	10 S77473	S77473 VNK-nontre
21	2274.5	85.2	1651	10 MMU05210	U05210 Mus musculu
22	2020.5	75.6	1713	6 AX644057	AX644057 Sequence
23	1592.5	60.4	246796	2 AC094643	AC094643 Rattus no
24	1592.5	59.6	247196	2 AC073822	AC073822 Mus muscu
25	1591.5	59.6	218817	2 AC114704	AC114704 Rattus no
26	1540.5	57.7	8212	6 AX306276	AX306276 Sequence
27	1540.5	57.7	8212	10 MMHXYLTJK	X83972 M. musculus
28	1501	56.2	43190	9 AC005777	AC005777 Homo sapi
29	1469.5	55.0	16389	6 AX644059	AX644059 Sequence
30	1269	47.5	738	6 AR053289	AR053289 Sequence
31	1253.5	46.9	1353	5 CHKSCKCA	M85039 Chicken src
32	1252	46.9	2292	10 BC018394	BC018394 Mus muscu
33	1252	46.9	2363	10 BC052006	BC052006 Mus muscu
34	1247.5	46.7	2187	6 AX330746	AX330746 Sequence
35	1247.5	46.7	2187	9 HSCSCKTK	X59932 Human mRNA
36	1247.5	46.7	2420	9 RPTYKI	X58631 Rat mRNA fo
37	1243.5	46.6	2125	10 MMU05247	U05247 Mus musculu
38	1227.5	46.0	1779	10 AY007162	AY007162 Homo sapi
39	1217.5	45.6	1682	5 AF052430	AF052430 Xenopus l
40	1217	45.6	1806	5 AK113892	AK113892 Clona Int
41	1057.5	37.6	2205	3 AB006558	AB006558 Ephydalia
42	1004.5	37.6	1733	3 AF067775	AF067775 Hydra vul
43	976.5	36.6	1566	3 AB096875	AB096875 Caenorhab
44	873.5	32.7	2504	3 AC073794_2	Continuation (3 of
45	838.5	31.4	110000	2	

RESULT 1

ALIGNMENTS

AR053292 AR053292 1942 bp DNA linear PAT 29-SEP-1999
LOCUS AR053292
DEFINITION Sequence 11 from patent US 5834208.
ACCESSION AR053292
VERSION AR053292.1 GI:5978154
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1942)
AUTHORS Sakano,S.
TITLE Tyrosine kinase
JOURNAL Patent: US 5834208-A 11 10-NOV-1998;
FEATURES Location/Qualifiers
source 1. 1942
BASE COUNT 365 a 614 c 652 g 311 t
ORIGIN
Alignment Scores:
Pred. No.: 1.93e-158 Length: 1942
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 gaps: 0
US-09-977-260-2 (1-507) x AR053292 (1-1942)
QY 1 MetAlaGlyArgGlySerLeuValSerProArgAlaPheHisGlyCysAspSerAlaGlu 20
Db 208 ATGGCGGGGAGAGCTCTCTGGTTCTCTGGGGGCAATTCAGAGCTGTATCTGCTGAG 267
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaThrPheIleProProValSerAla 40
Db 268 GAACCTCCCGGGGTGAGCCCCCGCTTCCTCGAGCCGGGCAACCCCTCCCGTCTAGCC 327
QY 41 ArgMetProThrArgArgTTPAlaProGlyThrGlnCysIleThrIleCysGluHisThr 60
Db 328 AGGATCCCAAGAGGGGCTGGGCGCCGGGCAACAGTATCACAATCCGAGCACACC 387
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
Db 388 CGGCCCAAGCAGAGGAGCTGGGCTTCGCAAGGGGAGCTGTCACCATCTGGAGGCC 447
QY 81 CysGluAsnLysSerTTPArgValLysHisHisThrSerGlyGlnGluLeu 100
Db 448 TGGGAGAACAAGAGCTGGTACCGCGTCAACACCAACCAAGTGAGAGGGGCTGCTG 507
QY 101 AlaIleGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db 508 GCAGCTGGGGGCTGGGGGAGCGGAGGCCCTCTCGCAGAACCCCAAGCTCACCTCATG 567
QY 121 ProThrPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
Db 568 CCTGGTTCACGCGAGAGATCTGGGCGCAGAGAGCTGTCCAGCAGCTGCCGCTCCGAG 627
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysVal 160
Db 628 GATGGGCTGTCTCTGGGGGAGTCCGGCGCCACCCCGGAGTACGTCCTGCTGCTG 687
QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
Db 688 AGCTTTGGCGCGACATCCACCTACCGGCTCTCCACCGCAGCGCCCACTCACAAATC 747
QY 181 AspGlyAlaValPhePheCysAsnLeuMetAspMetValGlnIleTyrSerLeuAspLys 200
Db 748 GATGAGGCGCTGTCTCTTCGCAACCTCATGAGCATGTGGAGGATTCACGCAAGACAG 807
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
Db 808 GGGGCTATCTGCACCAAGCTGTGTGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 867

QY 221 GluGluLeuAlaArgAlaGlyTTPLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
Db 868 GAGAGCTGGCCAGAGGGGCTGTACTGAACCTGCAGCATTTGACATTTGGAGGCACAG 927
QY 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyLysIleThrLeuGlyGlnValAla 260
Db 928 ATCGAGAGGAGAGGTTGGAGCTGTCTCAGGGTGAATACCTGGGGCAAAAGGTGCC 987
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
Db 988 GTGAGAAATATCAAGTGTATGTGACAGCCAGCCCTCTCGACAGAGAGCCGCTCATG 1047
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
Db 1048 ACGAAGATGCAACACGAGAACCTGTGCTCTCTGGGCTGATTCCTGCACAGAGGCTG 1107
QY 301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
Db 1108 TCAATGTCTATGGAGCAGTGTACCAAGGCAACCTGTGTGAATTTCTGCGGACCCGGGT 1167
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGlyMet 340
Db 1168 CGAGCCCTCGTGAACACCGCTCAGCTCTCAGATTTTCTGACAGTGGCCGAGGCGATG 1227
QY 341 GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
Db 1228 GAGTACCTGGAGAGCAAGACTTGTGACCGGAGACCTGTGCGCCCAACATCTGTGTC 1287
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
Db 1288 TCAGAGACCTGTGGCCCAAGTCAAGGACTTTGGCTGTGCGCAAAAGCCGAGCGAAAGGG 1347
QY 381 LeuAspSerSerArgLeuProValLysTTPThrAlaProGluAlaLeuLysHisGlyLys 400
Db 1348 CTAGAGTCAAGCCGGCTGCCCGTCAAGTGAGCGCGCCGACGCTGCAAAACAGGGAG 1407
QY 401 PheThrSerLysSerAspValTTPSerPheGlyValLeuLeuThrGluValPheSerTyr 420
Db 1408 TTCACCAAGCAAGTGGAGTGTGAGTTTGGGGTGTGCTCTCGGAGAGGCTTCTCATAT 1467
QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGlyValSerGluAlaValGluLysGly 440
Db 1468 GGAAGGCTCTGTAACCTTAAATGTCACTAAAGAGGTGTGCGAGCCGTGGAGAAAGGG 1527
QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
Db 1528 TACCGCATGAAACCCCGCGAGGCTGTCCAGGCCCGCTGCACGCTCATGAGACAGCTGC 1587
QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db 1588 TGGGAGGCAAGCCCGCCCGCCGCAACCTTCCGCAAACTGGCGAGAAAGCTGGCCCG 1647
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
Db 1648 GAGCTACGCAAGTCAAGTGTGCCCAAGCTTCGCTCAGAGGAGAGACCGCCAGGCTCCACC 1707
QY 501 SerProArgSerGlnGluPro 507
Db 1708 TCGCCCGGAAGCCAGAGGCC 1728
RESULT 2
LOCUS E10345 1942 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding human novel tyrosine kinase.
ACCESSION E10345
VERSION E10345.1 GI:22027175
KEYWORDS JP 1995313157-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1942)
AUTHORS Sakano,S.
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE
JOURNAL
COMMENT

NEW TYROSINE KINASE
Patent: JP 1995313157-A 1 05-DEC-1995;
ASAHII CHEM IND CO LTD
OS Homo sapiens (human)
PN JP 1995313157-A/1
PD 05-DEC-1995
PF 11-AUG-1994 JP 1994189444
PR 25-AUG-1993 JP 93P 210403, 29-MAR-1994 JP 94P 58553 PI
SAKANO SEIJI
PC C12N9/12, A61K38/45, C07K14/47, C12N15/09, (C12N9/12, C12R1.91), PC
(C12N15/09,
PC C12R1.91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1.1942
FT /organism='Homo sapiens'
FT /cell_type='megakaryoblast'
FT /cell_line='UT-7'
FT 5'UTR 1.207
FT CDS 208..1731
FT /product='human novel tyrosine kinase' FT
FT 3'UTR 1732..1942
FT CDS 331..1731
FT /product='human novel tyrosine kinase' FT
FT misc_RNA 349..540
FT /product='SH3 domain'
FT /product='SH2 domain'
FT misc_RNA 571..795
FT /product='SH2 domain'
FT 904..1641
FT /product='tyrosine kinase domain'.
FEATURES
source 1.1942
/organism='Homo sapiens'
/mol_type='genomic RNA'
/db_xref='taxon:9606'
BASE COUNT 365 a 614 c 652 g 311 t
ORIGIN
Alignment Scores:
Pred. No.: 1.93e-158 Length: 1942
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-977-260-2 (1-507) x E10345 (1-1942)
QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
DB ATGGCGGGGCGAGGCTCTCTGTTTCCTGCGGGGCGATTTCAGGGCTGATTCGCTAG 267
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla 40
DB GAACCTCCCGGGGTAGCCCGCTCTCTCCGAGCGTGCACCCCGCTCCGCTCAGCC 327
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
DB AGGATGCCACAGAGGCGCTGGGCCCGGCGACCAAGTGTATCACCAATGCGACACACC 387
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
DB CGGCCCAAGCCAGGGGAGCTGCTCCGCAAGGCGACGTCGTCACCATCTCGAGGCC 447
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluGlyLeuLeu 100
DB TGCAGAACACAGAGCTGTCAGCGGTCAAGACACACACAGTGTGACAGAGAGGCTGCTG 507
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120

DB GCAGCTGGGGCGCTCGGAGGAGGGAGCCCTCTCCGAGACCCCAACCTACGCTCATG 567
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
DB CCGTGGTTCACAGGAGATCTCGGCGCAGAGGCTGTCCAGACGCTCAGCTCCCGAG 627
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
DB GATGGGCTGTTCTCTGTCGAGAGTCCCGGCGACCCCGGAGTACGTCCTGCGCTG 687
QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
DB AGCTTGGCGGAGAGTCATCATCAGCGCGTCGACCGCGAGCGGCGGCTCATCATC 747
QY 181 AspGluAlaValPhePheCysAsnLeuMetLysPheValGluHisTrpSerLysAspLys 200
DB GATGAGGCCGTGTTCTCTGCAACCTCATGAGCATGTGAGCATTTACAGCAGAGCAG 807
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
DB GCGCTATCTGCACCAAGCTGTGAGACCAAGGAGAACACGGAGCCAAAGTCCGCGAG 867
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
DB GAGGAGCTGGCGAGGGCGGCTGTTACTGAACTGCAGCATTTGACATTGGAGCACAG 927
QY 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyIleTrpLeuGlyGlnLysValAla 260
DB ATCGAGAGGAGAGATTGTGAGCTGCTCTCGAGGTGAGTACCTGGGCGCAAAAGGTGCC 987
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLysAspGluThrAlaValMet 280
DB GTGAAGATATCAAGTGTGATGTGACACCCAGGCTTCTCGAGAGAGACGCGCTCATG 1047
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyAlaIleLeuHisGlnGlyLeu 300
DB ACGAAGATGCAACAGAGAACCTGTGCTCTCGGCGCGAGTCTGACACAGGCGCTG 1107
QY 301 TyrIleValMetGluHisValSerLysGlyLysLeuValAsnPheLeuArgTrpArgGly 320
DB TACATTGTCATGAGGACAGCTGAGCAGAGGCAACCTGTGATCTTCCGCGACCCGGGCT 1167
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
DB CGAGCCCTCGTGAACACCGCTCAGCTCTGTGAGTTTCTGCGACGTGGCGAGGCAATG 1227
QY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
DB GAGTACCTGGAGACAGAACCTGTGACACCGGAGCTGGCGCGCCGCAACATCTGTGTC 1287
QY 361 SerLysPheLeuValAlaLysValSerAspPheIleLeuAlaLysAlaGluArgGly 380
DB TCAAGAGACCTGTGGCCAAAGTGTGAGACTTGTGGCTGGCGAAAGCGGAGCGAAGGG 1347
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB CTAAGCTCAACCGCGTCCCGCTCAAGTGAAGCGCGCCGAGGCTCTCAAAACAGCGAAG 1407
QY 401 PheTrpSerLysSerAspValTrpSerPheGlyAlaLeuLeuTrpGluValPheSerTrp 420
DB TTCAACAGCAAGTGTGATGTGAGTTTGGGCTGCTGCTGGAGAGCTTCTCATAT 1467
QY 421 GlyValAlaProTrpProLysMetSerLeuLysGlyValSerGluAlaValGluLysGly 440
DB GGAGGGGCTCGTACCTTAATAATGTCACTAAAGAGGTGTGGAGGGCGGTGGAGAAAGGG 1527
QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
DB TACCGCATGGAACCCCGGAGGGCTGTCCAGGCGCCGTGACGTCCTCATAGACAGCTGC 1587
QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480

Db 1588 TGGGAGGACGAGCCCCCGGCGCACCCCTCCGCAAACTGGCCGAGAACTGGCCCG 1647
 QY 481 GlutLeuArgSerIaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 Db 1648 GAGCTACGACGATGTCAGAGTGGCCCGACGCTCGCTCTCAAGGAGAGAGCCGACGCTCCAC 1707
 QY 501 SerProArgSerGlnGluPro 507
 Db 1708 TCGCCCCGAGAGCCGAGAGCC 1728
 RESULT 3
 LOCUS HSHYTK 1968 bp mRNA linear PRI 10-OCT-1994
 DEFINITION H.sapiens HYL tyrosine kinase mRNA.
 ACCESSION X77278
 VERSION X77278.1 GI:471312
 KEYWORDS HYLTK gene; nonreceptor protein tyrosine kinase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Sakano, S., Iwama, A., Inazawa, J., Ariyama, T., Ohno, M. and Suda, T.
 TITLE Molecular cloning of a novel non-receptor tyrosine kinase, HYL
 JOURNAL (hematopoietic consensus tyrosine-lacking kinase)
 MEDLINE Oncogene 9 (4), 1155-1161 (1994)
 PUBMED 8134117
 REFERENCE 2 (bases 1 to 1968)
 AUTHORS Iwama, A.
 TITLE Direct Submission
 JOURNAL Submitted (14-JAN-1994) A. Iwama, Dept of Cell Differentiation,
 Inst of Mol Embryology & Genetics, Kumamoto University School of
 Medicine, 2-2-1 Honjo, Kumamoto 860, JAPAN
 FEATURES
 source
 1..1968
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="19p13"
 /tissue_type="hematopoietic cell"
 /clone_lib="UT-7 cDNA library"
 1..1968
 /gene="HMYTK"
 1..1968
 /gene="HMYTK"
 208..1731
 /gene="HMYTK"
 /product="HYL tyrosine kinase"
 /protein_id="CA54493.1"
 /db_xref="GI:557272"
 /translation="MAGRSLSVSMRAFHGCDSEELPRVSPFLRAMHPVARSMT
 RMAPDCTOCITCEHTRPKDGLAPKGDVITLACEKSKYRYKHHSOGDGLIA
 GALKREHLSADPKSLIMPFHGKISGDAVOLODPENGCLTVRESANHPGDIYLCV
 SEGRDVIHRYVLHROGHLTIDAEVFECNLMDEHYSKKGAICTKLVAPKRRHKTS
 AEEELARAGMLNLOHLLTGAQIGEGEFAVLQEGYIGQVAIKTKICDVTQAQFIDE
 TAVMTKMOHENVRLIGVILHOGLYVMHVGKGLNVLNFTGRALVTAQLOFSL
 HVAEGMEYLSKKLVHRDLAARNILVSEDLVAKVSPFGAKAKRGKLDSSRLPVKMTA
 PEAKHGKRTSKSDVMSRGVILMEVPSYRARYPKSKLSKESEAVKCGKRMPEPCP
 GPVHYLMSSCMLEAFARRPFFKRLAEKRLKLSAGAPASVSGDDADGSTSPSDP"
 BASE COUNT 391 a 614 c 652 g 311 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,966-158
 Score: 2671.00 Length: 1968
 Percent Similarity: 100.00% Matches: 507
 Best Local Similarity: 100.00% Conservative: 0
 Query Match: 100.00% Mismatches: 0
 DB: 9 Indels: 0
 Gaps: 0

US-09-977-260-2 (1-507) x HSHYTK (1-1968)
 QY 1 MetaIaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 Db 208 ATGGCGGGGCGAGGCTCTGTGGTTCTCTGGCGGCAATTCACGGCGTGTATTCGTGTAG 267
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
 Db 268 GAACCTCCCCGGGTGAGCCCCCGCTTCCTCCAGCTGGCACCCCCCTCCCTCAGGCC 327
 QY 41 ArgMetProThrArgArgTrpAlaProGlyTrpGlnCysIleThrCysGlnHisThr 60
 Db 328 AGATGTCACAGAGGGGCTGGGCGCCCGGACACCACTGATACCAAAATGCGAGCACACC 387
 QY 61 ArgProLysProGlyGluLeuAlaPheArgGlyGlyAspValValThrIleLeuGluAla 80
 Db 388 CGGCCCAACGACGAGGAGCTGGGCTTCGCCAAGAGGCGAGCTGTACATCTCGAGAGGCC 447
 QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluLeu 100
 Db 448 TGGGAGAACAGAGCTGTACCGCCGTCAGACACACACAGTGAGACAGAGGGCTGTCTG 507
 QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 Db 508 GCACTGGGGCGCTGGCGGAGCGGAGGCCCTCTCCGACAGACCCCAAGCTCAGCTCATG 567
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGluLeuProGlu 140
 Db 568 CCGTGTTCACAGGAGATCTCGGCGCAGAGAGCTGTCCACAGCTGCAGCTCCGAG 627
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaAlaGlnHisProGlyAspTrpValLeuGlyVal 160
 Db 628 GAGGGCTGTCTCTGTGGGAGGAGTCCGCGCCGACCCCGGAGCATCTCTGTGCGTG 687
 QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
 Db 688 AGCTTGGCGCGGACGCTCATCCACTCCGCTGCTACCGCGAGCGCCACCTCACAATC 747
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTrpSerLysAspLys 200
 Db 748 GATGAGCGCGGTGTTCTGTCAACCATGATGATGATGATGATGATGATGATGATGATG 807
 QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgGlyHisGlyLysSerLysGlu 220
 Db 808 GGGCTATCTGCACCAAGTGTGAGACCAAGGCAACGAGGCAAGTGGCGGAG 867
 QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGluAlaGln 240
 Db 868 GAGGAGCTGCCAGCGCGGCTGTACTGACMACTGACATTTGATGGAGACAG 927
 QY 241 IleGlyGluGlyLeuPheGlyAlaValLeuGlnGlyLysTrpLeuGlyGlnLysValAla 260
 Db 928 ATCGGAGAGGAGAGATTGTGAGCTGCTCGACAGGAGTACTGTGGCAAAAGTGCGCC 987
 QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
 Db 988 GTGAAGAATAATCAAGGTATGTGACAGCCAGGCGCTTCTGACAGAGCGCGCTGACG 1047
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnLysLeu 300
 Db 1048 ACGAAGATCAACACAGACAGACTGCTGCTCTCTGCGCTATCTCTGACCAAGGCGCTG 1107
 QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 Db 1108 TACATTGTGATGAGACGTGAGACAGGCAACGCTGTGACTTCTGCGGAGCCGCGGCT 1167
 QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet 340
 Db 1168 CGAGCCCTGTGTACACCCCTCAGCTCCCTGAGATTCTCTCAGCTGCGCAGGCGCAAG 1227
 QY 341 GluTrpLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 Db 1228 GAGTACTGTGAGAGCAAGAGCTTGTGACCGCGAGCTGTGCGCGCGGCAACATCTCTGCTC 1287

QY 361 SerGluAspLeuValAlaIalysValSerAspPheGlyLeuAlaIalysAlaGluArgLysGly 380
 Db 1288 TCACAGGACCTGTGGCCAGAGGTACGACCTTTGGCCGCAAGCGAGCGAGAGGG 1347
 QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 Db 1348 CTACACTCAGACCGCGCTGCCGTCAAGTGGACGGCGCCGAGGCTCTCAACACAGGGAG 1407
 QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerThr 420
 Db 1408 TTCACAGCAGAGTGTGATGTCTGGAGTTTGGGGTGTGCTGTGGAGAGCTTCTCATAT 1467
 QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGlyValSerGluAlaValGluLysGly 440
 Db 1468 GGAGGGGCTCCGTACCTAAATGTCACTGAAAGAGGTGTGGAGCGCGTGAAGAAGGGG 1527
 QY 441 TyrArgMetGluProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
 Db 1528 TACCGCATGGAACCCCGGAGGGCTGTCCAGGCCCGCCGACGTCCTCATAGCAGCTGC 1587
 QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
 Db 1588 TGGGAGGAGAGCCCGCGCGCGCACCTTCGCCAAGCTGCCGAGAACCTGGCCCGG 1647
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyLysAspAlaAspGlySerThr 500
 Db 1648 GAGCTACCGCAGTGCAGGTGGCCCGCCAGCTCCGTCTCAGGGCAGAGCGCGAGGCTCCACC 1707
 QY 501 SerProArgSerGluGluPro 507
 Db 1708 TCGCCCGCAGACGAGAGGCC 1728
 RESULT 4
 BDI76710
 LOCUS BDI76710 1997 bp DNA linear PAT 18-MAR-2003
 DEFINITION Method of examining allergic disease.
 ACCESSION BDI76710.1 GI:29122420
 VERSION WO 02075304-A/11.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1997)
 AUTHORS Sugita,Y., Heishi,M., Kagaya,S., Gunji,S. and Tsujimoto,G.
 TITLE Method of examining allergic disease
 JOURNAL Patent: WO 02075304-A 11 26-SEP-2002;
 GENEX RESEARCH INC., JAPAN AS REPRESENTED BY GENERAL DIRECTOR OF NATIONAL CHILDREN'S HOSPITAL, HIDEYO YONEKURA, YASUHIKO YANAMOTO, SHIGERU SAKURAI, TAKIO MATANBE, YUKI SUGITA, MASAYUKI HEISHI, SHINJI KAGAYA, SHIGEMICHI GUNJI, GOZO TSUJIMOTO
 COMMENT OS Homo sapiens (human)
 PN WO 02075304-A/11
 PD 26-SEP-2002
 PR 01-MAR-2002 WO 2002JP001916
 PI 21-MAR-2001 JP 01P 081028
 PI YUKI SUGITA, MASAYUKI HEISHI, SHINJI KAGAYA, SHIGEMICHI GUNJI, PI GOZO TSUJIMOTO
 PC GO1N33/15,GO1N33/50,C12N15/09,C12Q1/02,C12Q1/68,A61K45/00, PC A61K31/711,
 PC A61K39/395,A61P37/08,A61K48/00,A01K67/027
 CC Method of examining allergic disease
 FH key Location/Qualifiers
 FT CDS Location/Qualifiers
 FEATURES
 source 1..1997
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 372 a 631 c 676 g 318 t
 ORIGIN

Alignment Scores:

Pred. No.: 2e-158 Length: 1997
 Score: 2671.00 Matches: 507
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-977-260-2 (1-507) x BDI76710 (1-1997)

QY 1 MetaGluArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 Db 266 ATGGCGGGGAGAGCTCTGTGTTCTCTGGGGGACATTTCACGGCTGTGATTCGTGGAG 325
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla 40
 Db 326 GAACTTCCCGGGGAGACCCCGGCTTCCGAGCCGAGACCCCGCTCCGTCAGCC 385
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
 Db 386 AGGATGCCAACAGAGCGCTGGGCCCGGGCACCCAGTGTATACCAATGCGAGACACACC 445
 QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
 Db 446 CGCCCAAGCCAGGGGAGCTGCTTCCGCAAGGGGACGTCGACCATCTGGAGGCC 505
 QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGluGlyLeu 100
 Db 506 TGCAGAAACAAGACTGGTACCGCGTCAAGCACACACCAATGAGACAGAGGGCTGTG 565
 QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 Db 566 GCAGCTGGGGCGCTCGGGAGAGGGAGGCCCTCTCCGAGACCCCAAGCTCAGCTCATG 625
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuProProGlu 140
 Db 626 CCGTGTTCACCGGAGAGATCTCGGGCAGAGGCTGTCCAGCAGCTGCACCTCCGAG 685
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
 Db 686 GATGGCGCTTCGTGGTGGGAGTCCGGCGGCACCCCGGCGACTAGCTCTGGGTG 745
 QY 161 SerPheGlyArgAspValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
 Db 746 AGCTTGGCCGCGACGTCATCTACCTACCGCTGTGCACCGCGACGGCACCCTCAATC 805
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTrpSerLysAspLys 200
 Db 806 GATGAGCGCGTGTCTCTGCAACCTCATGACATGTGGAGCATTTACAGCAGACAGAG 865
 QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 Db 866 GGCGCTATCTCCACAGAGCTGTGAGACCAAGGGAGGAACCGGAGCAACCAAGTCCGGAG 925
 QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 Db 926 GAGAGCTGGCGCAGGGCGGGCTGTGTTACTGAACCTGCAGCATTTGACATTGGAGCAG 985
 QY 241 IleGlyGluGlyLeuPheGlyAlaValLeuGlnGlnGlyLysTrpLeuGlnLysValAla 260
 Db 986 ATCGGAGAGGAGAGATTGGAGCTGCTCTGCAGGCTGAGTACCTGGGGCAAAAGGTGCC 1045
 QY 261 ValLysAsnIleLysCysAspValTrpThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
 Db 1046 GTGAAGATATCAAGTGTGATGTGACACCCAGGCTTCTGTGAGAGAGACGGCTGTCTG 1105
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyAlaIleLeuHisGlnGlyLeu 300
 Db 1106 AGCAAGATGCAACACAGAACCTGGTGTCTCTGGGCGTGATCTGCACACAGAGGGCTG 1165
 QY 301 TyrIleValMetGluHisValSerLysGlyLysLeuValAsnPheLeuArgThrArgGly 320
 Db 1166 TACATTGTTCATGGAGCAGTGTGAGCAAGGGCAACCTGTGTGAATTTCTGCGAGACCGGGGT 1225

QY	321	ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet	340
Db	1226	CGAGCCCTCGGAAACACCCTCAGCTCCTCGAGTTTCTCTGCACGTCGGCCGAGGATG	1285
QY	341	GluTyrLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal	360
Db	1266	GAGTACCTCGGAGAGCAAGAACCTTGTGACCGCAGCCTGGCCGCCGCAATCCTGTGTC	1345
QY	361	SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly	380
Db	1346	TCAGAGGACCTGTGTGGCCCAAGCTCAGCAGCATTTGGCCCTGGCCCAAGCCGAGGAGGG	1405
QY	381	LeuAspSerSerArgLeuProValLysTyrPheAlaProGluAlaLeuLysHisGlyLys	400
Db	1406	CTAGACTCAACCCGGCTCTCCCGTCAAGAGGAGGGGCCGAGGCTCTCAACACGGGAG	1465
QY	401	PheThrSerLysSerAspValTyrPheSerPheGlyValLeuLeuTyrPgluValPheSerTyr	420
Db	1466	TTCAACACCAAGTCGAGTGTCTGAGTTTGGGGTGGCTGTGGGAGGTTCTTCTCATAT	1525
QY	421	GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValAlaGluLysGly	440
Db	1526	GGAGCGGCTCCGTACCCCTAAATGTCATCTGAAGAGGTGTGGAGGGCCGTGAGAAAGGG	1585
QY	441	TyrTrpMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys	460
Db	1566	TACCGCATGGAACCCCGCAGGGCTGTCAAGGCCCGCCGTCACGTCTCATAGACGTCC	1645
QY	461	TyrPgluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg	480
Db	1646	TGGAGGCGAGAGCCCGCCCGCCGCCACCTTCCGCAACTGGCCGAGAAAGCTGGCCGG	1705
QY	481	GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGluAspAlaAspGlySerThr	500
Db	1706	GAGCTACGCAATGAGAGGTGCCCCAGCCCTCCTCAAGGCGAGAGCGCAGCGCTCCACC	1765
QY	501	SerProArgSerGlnGluPro 507	
Db	1766	TCGCCCGAAGCCAGGAGCC 1786	
RESULT 5			
LOCUS	ARI179662	2000 bp	DNA
DEFINITION	Sequence 1 from patent US 6326469.		linear
ACCESSION	ARI179662		PAT 20-APR-2002
VERSION	ARI179662.1		
KEYWORDS	GI:20221217		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 2000)		
TITLE	Ullrich,A., Gishizky,M. and Sures,I.		
JOURNAL	Megakaryocytic protein tyrosine kinases		
FEATURES	Patent: US 6326469-A 1 04-DEC-2001;		
	Location/Qualifiers		
source	1..2000		
BASE COUNT	375 a 632 c 673 g 320 t		
ORIGIN	"/organism="unknown"		
Alignment Scores:			
Pred. No.:	2e-158	Length:	2000
Score:	2671.00	Matches:	507
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
	6	Gaps:	0
US-09-977-260-2 (1-507) x ARI179662 (1-2000)			
QY	1	MetaIaGlyArgGlySerLeuValSerTyrArgAlaPheHisGlyCysAspSerAlaGlu	20

Dd	258	ATGGCGGGGCGAGGCTCTCTGGTTTCTGGCGGGGCAATTCACGGCTGTGATTCCTGTGAG	317
Qy	21	GIuIeuProAqyValSerProArghPheIuArGlaATrPhISProProIValSerAla	40
Dd	318	GAACCTTCCCCGGGTAGGCCCCCCGCTTCTCCGAGGCTGGACACCCCTTCCTGTACAGC	377
Qy	41	ArgMetProThrArgArgTyrPrlaProGlyThrGlnCysIleThrLysCysAluHisThr	60
Dd	378	AGAGATGCCAAGAGAGCGCTGGGCCCCCGGAGACCCAGTGTATACCAAAATGCGAGCACACC	437
Qy	61	ArgProLysProGlnGlyIuIeuAlaPheArgLysGlyAspArgValThrIleIeuAluAla	80
Dd	438	CGCCCCAAGCCAGGGGAGCTGGCTCTCCGCAAGGGCGAGCGTGTACCAATCTCTGAAGGCC	497
Qy	81	CysGluAsnLysSerTyrPrlArgValLysHisHisThrSerGlnGlnGlyIuIeuLeu	100
Dd	498	TGCGAGAACAAAGAGCTGGTACCGGCTCAAGACACACACAGTGGAGACAGAGGGGCTGCTG	557
Qy	101	AlaAlaGlyAlaLeuAlaArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet	120
Dd	558	GCAGCTGGGGCGGCTCGGGAGCGGAGGCGCCCTCTCCGAGACCCCAACTCGACCTCATG	617
Qy	121	ProTrpPheHisGlyLysIleSerLysGlnGlnAlaValGlnGlnLeuGlnProProGlu	140
Dd	618	CCGTGTGTCCACGGGAAATATCTCGGGCCAGAGGGGTGTCCAGACACTGCATCTCCGAG	677
Qy	141	AspGlyLeuPheLeuValArgIuSerAlaArgHisProGlyAspTyrValLeuCysVal	160
Dd	678	GATGGGCTGTTCCTGTGGTGGGAGTCCGCCGCCACCCCGCGAGTACGTCTGTGGCTG	737
Qy	161	SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle	180
Dd	738	AGCTTTGGCCCGAGCATCATCACTAACCGCGTGTGCACCGGAGCGGCACCTCACAATC	797
Qy	181	AspGluAlaValPhePheCysAsnLeuMetLysMetValGluHisTyrSerLysAspLys	200
Dd	798	GATGAGGCCGAGTGTCTTCTGTGCACCTCATGGACAGTGGAGCATTAAGCAAGGACAG	857
Qy	201	GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu	220
Dd	858	GGCGCTATCTGCACCAAGCTGGTGGAGACCAAGCGGAAACACGGGACCAAGTGGCCGAG	917
Qy	221	GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln	240
Dd	918	GAGGAGCTGGCCAGGCGCGGCTGGTTACTGAACCTCGACAGCATTTGACATTTGGAGCACAG	977
Qy	241	IleGlyGlnGlyGlnPheGlyAlaAlaValLeuGlnGlyGlyTyrLeuGlyGlnLysValAla	260
Dd	978	ATCGGAGAGGAGAGGTTTGGAGCTGTCTCGAGGGTGTGATCACTGGGCGCAAAAGGTGGCC	1037
Qy	261	ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMet	280
Dd	1038	GTTGAAGATATCAAGTGTGATGTGACAGCCAGGCTTCTGTGAGAGAACGGGCTCATG	1097
Qy	281	ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu	300
Dd	1098	ACGAAGATGCAACACGAGAACCTGGTGGCTCTCGGGCGTATCTGCACAGGGGCTG	1157
Qy	301	TyrIleValIleMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuAlaThrArgGly	320
Dd	1158	TACATTGTTCAGAGGACAGCTGAGCAAAAGGCAACCTGTGTGAATCTTCTCGGACCCGGGT	1217
Qy	321	ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMet	340
Dd	1218	CGAGCCCTCGGAACACCGCTCAGCTCTCGCAATTTTCTTGTGCACGTGTGGCGAGGCAATG	1277
Qy	341	GluTyrLeuGlnIuSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal	360
Dd	1278	GAGTACCTGGAGACAGAAACCTTGTGCACCGCGACCTGGGCGCGCAACATCTGTGTC	1337
Qy	361	SerLysAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly	380
Dd	1338	TCAAGAGCACTGTGTGGCCAAAGGTACAGCACTTTGGCTGGCCAAAGCCGACCGGAAAGGG	1397

|||||
Db 1120 ATCGAGAGGAGAGATTGGAGCTGTCTCGACAGGAGTACCTGGGGCAAAAGGTGGCC 1179
Qy 261 VallysantileysCysaspvalThrAlaGlnAlaLeuaspGluThrAlaValMet 280
Db 1180 GTGAAGATATCAAGTGTGATGTACAGCCAGGCTTCTGTGACGAGAGCGCGTCAATG 1239
Qy 281 ThrLysmetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
Db 1240 ACGAAGATGCACACGAGAACCTGGTCTCTCTGGGCGTGTATCTTCACACAGGCGCTG 1299
Qy 301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPhleuAspArgThrArgGly 320
Db 1300 TACATGTCTAGTACGACAGTACGACAGGCAACCTGGTGAACCTTTCGGGACCGCGGGGT 1359
Qy 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPhnSerLeuHisValAlaGluGlyMet 340
Db 1360 CGAGCCCTCTGGAACACCGCTCACCTCTCTGACGTTTCTCTGACGTCGCGCGAGGCGCATG 1419
Qy 341 GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
Db 1420 GAGTACCTGGAGACAGAGAGCTGTGACCGGACCTGGCGCGCCGACATCTCTGTC 1479
Qy 361 SerGluAspLeuValAlaLysValSerAspPhleGlyLeuAlaLysAlaGluArgLysGly 380
Db 1480 TCAGAGACCTGTGGCCAGGCTCAGGACCTTGTGCTGGCCAAAGCCGAGCGAGAGGGG 1539
Qy 381 LeuAspSerSerArgLeuProValLysThrPhrAlaProGluAlaLeuLysHisGlyLys 400
Db 1540 CTAGACTCAAGCCGCGCTGCTCAAGTGAAGGAGCGCGCGGCTCTCAACACGGAAG 1599
Qy 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
Db 1600 TTCACACGACAGTGTGATGCTGTGAGTTTGGGGTGTGCTGTGGAGGTCTTCATAT 1659
Qy 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
Db 1660 GGAGGGGCTCGTACCTAAATGTCACTGAAAGAGTGTGAGGCGCGTGGAGAGAGGG 1719
Qy 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
Db 1720 TACCGCATGGAACCCCGGAGGCTGTCCAGGCGCCGCTGCACCTCATAGAGAGCTGC 1779
Qy 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db 1780 TGGGAGGACAGAGCCCGCGCGGCGCACCTTCGCAACCTGGCGGAGAGCTGGCCGG 1839
Qy 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyLysAspAlaAspGlySerThr 500
Db 1840 GAGCTACGACAGTGCAGGTGCGCCAGCTCCGTCACAGGCGAGAGCGCGAGCTCCACC 1899
Qy 501 SerProArgSerGlnGluPro 507
Db 1900 TCGCCCGGAGCCAGAGAGCC 1920

RESULT 8
BC003109
LOCUS
DEFINITION
Homo sapiens, megakaryocyte-associated tyrosine kinase, clone
MGC:2101 IMAGE:3357648, mRNA, complete cds.
ACCESSION
BC003109
VERSION
BC003109.1 GI:13111882
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 2180)
AUTHORS
Strausberg, R.
TITLE
Direct Submission
JOURNAL
Submitted (13-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590.

REMARK
COMMENT
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@gscc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Iliia Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Matsaja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

FEATURES
source
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: e Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4505108.
location/Qualifiers
1. 2180
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocustID:4145"
/db_xref="taxon:9606"
/clone="MGC:2101 IMAGE:3357648"
/tissue_type="Eye, retinoblastoma"
/clone_id="NIH_MGC_16"
/lab_host="DH10B-R"
/note="vector: pOTB7"
400. 1923
/codon_start=1
/product="megakaryocyte-associated tyrosine kinase"
/protein_id="AAH03109.1"
/db_xref="GI:13111883"
/translation="MAGRSGLVSRARFHCDSAEELPRVSPRELRAMHPPVSRMPT
RRAPGTCITKEHTRPKPELAFKRGDVVTILEACENKSWRVKHNHTSGOGLLA
GALREKALSDADPLSLMPWFHGKISQEAIVOOLOPEDELFLVRESARRHGDVLY
SGFQDVHYHVLHRDHGLTIDEAFPCNLMDVHYSKDGAICTKLVRRKRKGTGS
AEEELRAGLNLNQLTLGAOIGEGEFGVILGEGYGVAVKNTICDVTAAFLDE
TYAVTKQKHENLVRLCYITLHGCIYIYMEVSKSNLYNPLRTKRALVYRQLQSPSL
HVAEGMEYLESKLVHRDLARNILVSEDIVAVKSDGLKAKERKGLDSSRLPVKWA
PEALKHGRFTSKSDVMSFGVLLVGRAPYPMKMLKESVSEAVENGYMEPEPGP
GPVHLVMSCEWSEAPRRPFRKLAETLARELSAGAPASVSGODADGSPSPSOBP"

BASE COUNT
436 a 689 c 724 g 331 t
ORIGIN

Alignment Scores:
Pred. No.: 2,21e-158 Length: 2180
Score: 2671.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-977-260-2 (1-507) x BC003109 (1-2180).

Qy 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaIleHisGlyCysAspSerAlaGlu 20
Db 400 ATGGCGGGGCGAGGCTCTGTGGTTCTGTGGGGGCAATTCACGGCTGTGATCTGCGAG 459
Qy 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
Db 460 GAATCTCCCGGGGAGACCCCGCTCTCGAGCCGCGACACCCCGCTCCCGTCAAGCC 519
Qy 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
Db 520 AGGATGCCAACGAGGCGCTGGGCGCGGACACCAAGTATACCAATGCGAGACACAC 579

QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
 DB 580 CCCCCAACGAGGAGGAGCTGGCTCCGCAAGGGGAGCGTGTCACATCTCGAGAGGCC 639
 QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGluLeu 100
 DB 640 TCGGAACACAGAGCTGTACCGCTCAAGCACACACAGTGAGAGAGGGCTGTG 699
 QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 DB 700 GAGAGCTGGGGCTGGGGAGGAGGAGGCTCTCCGAGACCCCAAGCTCAGCCCTCAG 759
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProGlu 140
 DB 760 CCGTGGTTCACAGGAGATCTCGGGCCAGAGGCTGTCCAGACGCTGCAGCTCCCGAG 819
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal 160
 DB 820 GATGGCTGTCTCTGGTGGAGATCCGGCCGCCACCCGCGACTACTCTCTGTCGTG 879
 QY 161 SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
 DB 880 AGCTTTGGCCGGAGCTCATCCACTACCGCTGCTGCACCGGAGCGCCACTCACAATC 939
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLysAspLys 200
 DB 940 GATGAGGCGCTGTCTTCTGCAACCTCATGACATGTGAGCATTCACAGCAAGCAAG 999
 QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgGlyHisGlyThrLysSerAlaGlu 220
 DB 1000 GCGGCTATCTGCACCAAGCTGTGAGACCAAGCGGAAACAGGGGACCAAGTCGGCGGAG 1059
 QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlnAlaGln 240
 DB 1060 GAGGAGCTGGCCAGGGCGGGCTGTTACGAACTGTGCACATTGTGACATTGGAGACAG 1119
 QY 241 IleGlyGluGlyLeuPheGlyAlaValLeuGlnGlyGlyTyrLeuGlyGlnLysValAla 260
 DB 1120 ATCGGAGAGGAGATTTGTGAGCTGTCTCGAGGGGAGATACCTGGGGCAAAAGGTGGCC 1179
 QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
 DB 1180 GTCAGAAATATATCAAGTGTATGTGACAGCCAGGCTCTCTGACAGAGCGGCTCAAG 1239
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuGlyValIleLeuHisGlnGlyLeu 300
 DB 1240 ACGAAGATGCACACAGAACTGTGCTGTCTCGGGGATCTGTGACACAGGGGCTG 1299
 QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 DB 1300 TACATTGTATGAGAGACGTGAGCAAGGGCAACCTGTGTAACCTTTCTGCGGAGCCCGGCT 1359
 QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnIlePheSerLeuHisValIleAlaGlnGlyMet 340
 DB 1360 CAGAGCTCTGTAAACCCGCTCAGCTCTGCACTTTCTCTGACAGTGGCCGAGGGCAAG 1419
 QY 341 GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 DB 1420 GAGTACCTGAGAGCAAGAACTGTGTGACCGCGAGCTGGCGCGCAAACTCCGTGTC 1479
 QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
 DB 1480 TCAGAGGAGCTGTGGCCCAAGGTCAAGCGACTTTGGGCTGGCCAAAGCCAGCGGAGAGGG 1539
 QY 381 LeuAspSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 DB 1540 CTAGACTCAAGCCGCTGCCCTCAAGTGTGAGCGGCGCGGAGCTCTCAAAACAGGGAGG 1599
 QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
 DB 1600 TTCACACAGCAAGTGGATGTCTGTGAGTTTGGGGTGTCTCTGGGAGAGTCTTTCATAT 1659

QY 421 GlyArgAlaProTyrProLysMetSerLeuGlyGluValSerGluAlaValGluLysGly 440
 DB 1660 GAGAGGAGCTCGCTACCCCTAAATGTCAAGAAAGAGTGTCCGAGGCGCTGAGAAAGGGG 1719
 QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerGlyCys 460
 DB 1720 TACCGATGAAACCCCGGAGGAGCTGTCCAGGCGCCGTCACGTCTCTCATGAGCAAGCTTC 1779
 QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
 DB 1780 TGGAGGAGCAGACCCCGCGGCGGCGCACCTTCGCAAACTGGCGAGAAAGCTGGCGCGG 1839
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspLysThr 500
 DB 1840 GAGCTACGACGATGAGTGTCCCGAGCTCTGCTCAGGAGGAGGAGCGCGAGCTTCACAC 1899
 QY 501 SerProArgSerGlnGluPro 507
 DB 1900 TCGCCCGGAGGAGGAGGCC 1920
 RESULT 9
 AR053291 1521 bp DNA linear PAT 29-SRP-1999
 LOCUS AR053291
 DEFINITION Sequence 10 from patent US 5834208.
 ACCESSION AR053291
 VERSION AR053291.1 GI:5978153
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 1521)
 AUTHORS Sakano, S.
 TITLE Tyrosine kinase
 JOURNAL Patent: US 5834208-A 10 10-NOV-1998;
 FEATURES
 source location/Qualifiers
 1..1521
 BASE COUNT 295 a 470 c 507 g 249 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5, 32e-158 Length: 1521
 Score: 2662.00 Matches: 506
 Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 99.66% Indels: 0
 DB: 6 Gaps: 0
 US-09-977-260-2 (1-507) x AR053291 (1-1521)
 QY 1 MetalaglyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 DB 1 ATGGGGGGGAGGCTCTGTGTTCTCTGCGGGGATTCACGGCTGTGATCTGCTGAG 60
 QY 21 GluLeuProArgValSerProArgPheLeuAlaGlnAlaTrpHisProProValSerAla 40
 DB 61 GAACCTCCCGGGGTAGGCCCCCGCTTCCTCCAGGCTGGACACCCCTCCGCTCAGGC 120
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnHisThr 60
 DB 121 AGGATGCCACAGAGGCGCTGGGGCCCCGGGACCCAGCTGTATATCCAAATGCAAGACAC 180
 QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAla 80
 DB 181 CCGCCCAAGCCAGGGAGAGCTGTCCGCAAGGGGAGAGCTGTCCACCATCTCGAGGCC 240
 QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGluLeu 100
 DB 241 TCGGAGAACACAGAGTGTGTACCGCTCAAGCACACACAGTGTGACAGAGGGGCTGTG 300
 QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 DB 301 GCAAGCTGGGGCGCTCGCGGAGGAGGCGGAGGCTCTCCGCAAGACCCCAAGCTCAGCTCATG 360

QY	121	ProtrpPheHISGLYSYLIISeSeriGLInGluValaIngInGluInProProGlu	140
Db	361	CCGIGGTTCACAGGGAAAGATCTCGGGCCAGGAGGCTGTCAGCAGCTGCACCTCCCGAG	420
QY	141	AspGlyLeuPheLeuValaArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal	160
Db	421	GATGGGGCTGTTCTGTGTGGGAGATCCGGCGGCCACCGCGCGACTACGTCTGTGGCTG	480
QY	161	SerPheGlyArgAspValIIeHisTyrArgValLeuHisArgAspGlyHisLeuThrIle	180
Db	481	AGCTTTGGCCCCGAGCTCATCTACCTACCGCTGGAGCCCGCAGCGGCACTCAATC	540
QY	181	AspGluValaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys	200
Db	541	GATAGAGCGCGTCTTCTGTGCACCTCATGGACATGTGGAGCATATACGACGAGGACAG	600
QY	201	GlyAlaIleCysThrLysLeuValaArgProLysArgLysHisGlyThrLysSerAlaGlu	220
Db	601	GGCCCTATTCTCTACCAAGCTGGTAGACCAAAAGCGGAAACAGGGAGCCAAATCGGCCAG	660
QY	221	GluGluLeuValaArgAlaGlyTyrProLeuLeuAsnLeuGlnHisIleuThrLeuGlyAlaGln	240
Db	661	GAGAGAGCTGGCCAGAGGGGGGGCTGGTTACTGTGAACCTGCAGCATTTGTACATTGGAGCAG	720
QY	241	IleGlyGluGlyGluPheGlyAlaValaLeuGlnGlyGluTyrLeuGlyGlnLysValaIa	260
Db	721	ATCGAGAGGGAGATTGGAGCTGTCTCTGCAGGGGTGATGACTCGGGCCAAAGGTGGCC	780
QY	261	ValLysAsnIleLysCysAspValIThrAlaGlnAlaPheLeuAspGluThrAlaValMet	280
Db	781	GTTGAAGATATCAAGTGTGATGTGACACCCAGGCGCTTCTCGAAGACGACGGCGTCAATG	840
QY	281	ThrLysMetGlnHisGluAsnLeuValaArgLeuGluGlyValIIeLeuHisGlnGlyLeu	300
Db	841	ACGAAGATGCACACGAGAACCTGGTGGCTTCTCGGGCGTGAATCTGCACAGGGGCTG	900
QY	301	TyrIleValMetGluHisValaSerLysGlyAsnLeuValaAsnPheLeuArgIThrArgGly	320
Db	901	TACATTGTCAAGGAGACACTGACGAAAGGCAACTGGTGAATTTCTCGAGACCGGGGCT	960
QY	321	ArgAlaLeuValaAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValaIaGluGlyMet	340
Db	961	CGACCCCTCGGAACACCGCTCACCTCTCTGCAATTTCTTGCAAGCTGGCCGAGGGCATG	1020
QY	341	GluTyrLeuGluGluSerLysLysLeuValaHisArgAspLeuAlaAlaArgAsnIleLeuVal	360
Db	1021	GAGTACCTCGAGACGACGAAGACTTGTGCACCGCACCTGGCCCGCCGACACATCTGGTC	1080
QY	361	SerGluAspLeuValaAlaLysValaSerAspPheGlyLeuAlaLysAlaGluArgLysGly	380
Db	1081	TCAAGAGGACCTGGGGCCAAAGTCAAGCACATTTGGCCCGGCCAAAGCCGACGCGAAGGG	1140
QY	381	LeuAspSerSerArgLeuProValIleTyrPThrAlaProGluAlaLeuLysHisGlyLys	400
Db	1141	CTAACACTCAACCGCGGCTCCCGTCAAGTGAAGGAGCGGCCGAGGCTCTAAACACGGAGAG	1200
QY	401	PheThrSerLysSerAspValITrpSerPheGlyValLeuLeuTyrPgluValaPheSerTyr	420
Db	1201	TTTCAACAGCAAGTGGATGTCTGAGATTTTGGGGTGTGCTCTGGAGAGTCTTCTCATAT	1260
QY	421	GlyTyrAlaPheTyrProLysMetSerLeuLysGluValaSerGluAlaValaGluGlyGly	440
Db	1261	GGAGGGGCTCGGTACCTTAATAATGTCACTGAAGAGGTGTGGAGGCCGTGGAGAGGGGG	1320
QY	441	TyrArgMetGluProProGluGlyCysProGlyProValaHisValLeuMetSerSerCys	460
Db	1321	TACCGCATGAGAACCCCGCAGGGGCTGTCAAGGCCCGCGCAAGTCTCATGAGCAGTGC	1380
QY	461	TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg	480
Db	1381	TGGAGAGGAGAGCCCGCCCGCGGCACACTTCCGCAAACTGGCCGGAAGACTGGCCGG	1440

QY	481	GlueuhtgyserractagialyAlaProLaserValSerGlyGlnAspAlaAGGlySerThr	500
Db	1441	GAGCTACGACCAATGAGGAGGCCCAAGCTCTGCTCAGGAGCAGACGCCGAGGCTCACCC	1500
QY	501	SerProArySerGlnGluPro	507
Db	1501	TCGCCCCGACCAAGCAGAGCCC	1521
RESULT 10			
LOCUS	S71669	1744 bp	mRNA linear PRI 24-OCT-1994
DEFINITION	1sk-leukocyte carboxyl-terminal src kinase related gene [human,		
ACCESSION	S71669		
VERSION	S71669.1	GI:559593	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	McVicar,D.W., Lai,B.K., Lloyd,A., Kawamura,M., Chen,Y.Q., Zhang,X.,		
TITLE	1 (bases 1 to 1744)		
JOURNAL	Staples,J.E., Ortaldo,J.R. and O'Shea,J.J.		
MEDLINE	Molecular cloning of 1sk, a carboxyl-terminal src kinase (csk)		
PUBMED	related gene, expressed in leukocytes		
REMARK	Oncogene 9 (7), 2037-2044 (1994)		
FEATURES	GenBank staff at the National Library of Medicine created this		
Source	entry [NCBI g1bbseq 151068] from the original journal article.		
gene	This sequence comes from Fig. 1.		
CDS	Location/Qualifiers		
	1..1744		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	1..1744		
	/gene="1sk"		
	/note="leukocyte carboxyl-terminal src kinase related		
	gene"		
	125..1522		
	/gene="1sk"		
	/note="leukocyte carboxyl-terminal src kinase related		
	gene; this sequence comes from Fig. 1"		
	/codon_start=1		
	/protein_id="AAB30995.1"		
	/db_xref="GI:559594"		
	/translation="MPTRRMAPGTCITKCEHTPRPKGELAFRKGDVVTILACENKSS		
	MYRKHTNNGSGEGLLAGALREERLASDPRKLSLMPFHGKISGEAVOOLQDPEDLL		
	FLVSGHSGPDGYLVCTGFGDVIHYVYLHDDGLITIDENFPCNLMDMNYHSKDKG		
	ALICRLVLRPKRKHGKTSABEELAAAGNLINQHTLTAAQISGEFGAVIQEYIGQKC		
	VKNRLCDVTAQAFIDETAVMTKQHEHENVRLIGVILDQGLTYVHEHSGKLVNFLT		
	RGRLALNVTADQLDSLVHREGMELSKKLVNRDLAARNILVSEELAVSDFGLAFA		
	ERKQIDSSRLPVPKWTAPDELKHGFTSKSDVMSFGVILMEVTSGLAPYPMKSLKVS		
	EAVKRYRMPEEGCGPGFVHVMSSCWEAEAPARRPRKLAELKRLRSAGAPASVS		
	GQDDGSGSPRSQEP".		
BASE COUNT	354 a 534 c 582 g 274 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.62e-154	Length:	1744
Score:	2607.50	Matches:	497
Percent Similarity:	98.62%	Conservative:	6
Best Local Similarity:	98.22%	Mismatches:	2
Query Match:	97.62%	Indels:	1
DB:	9	Gaps:	1
QY	2	AlaGlyAryGlySerLeuValSerTTPArgAlaPheHISGlyCysAspSerAlaGluGlu	21
Db	5	GCGGGGCGAGGCTCTCTGTGGTTCCCTGCGGGGCAATTCACGGCTGTGATTCGTCGAGAA	64

QY 22 LeuProArGValSerProArGpHleuArGAlaTrpHisProProProValSerAlaArg 41
 |||
 Db 65 CTTCGCCGGGTGAAGCCCCCTTCCTCCGAGCCGGGAGCCCCCTCCCGCTCCAGCCAGG 124
 QY 42 MePProthArGArGTrpAlaProGlyThrGlnCysIleThrIlyScySGluHisThrArg 61
 |||
 Db 125 ATGCCAACGAGGCCCTGGCCCGGCGACCCAGTATCAACAAATCGAGCACACC 184
 QY 62 ProIlySerProGlyGluLeuAlaPheArGlySGlyAspValAlaThrIleLeuGluAlaCys 81
 |||
 Db 185 CCCAAGCCAGGGAGCTGGCCCTCCGCAAGGGGACCTGGTACACCATCTCGAGGGCTGC 244
 QY 82 GluAsnIlySerTrpIlyTrpArGValLysHisIleThrSerGlyGlnGluIlyLeuAla 101
 |||
 Db 245 GAGAACCAAGAGCTGGTACCCGCTCAACACACACCCAGTGTGACAGGGGGCTGCTGCA 304
 QY 102 AlaGlyAlaLeuArGValArGValAlaLeuSerAlaAspProIlySerLeuSerLeuMetPro 121
 |||
 Db 305 GCTGGGGCGCTCCGGGAGCGGGAGGCCCTCTCCGCAAGCCCAAGCTCAGCTCATGCGC 364
 QY 122 TrpPheHisGlyIlySerGlyGlnGluAlaValGlnIleuGlnProProGluAsp 141
 |||
 Db 365 TGGTCCAGCGAAGATCTCGGGCGAGAGGCTGTCCAGCAGCTGCAGCTCCCGAGAT 424
 QY 142 GlyLeuPheLeuValArGValSerAlaArgHisProGlyAspIlyValLeuCysValSer 161
 |||
 Db 425 GGGCTGTCTCTGCGGGAGCTCGGGCGGCCCGGCGACCTACGCTCTGTGCTGACG 484
 QY 162 PheGlyArGAspValIleHisTrpArGValLeuHisArGAspGlyHisLeuThrIleAsp 181
 |||
 Db 485 TTGGGGCGGACGTCATCCATCCGCTCTGACCGGCGACCGCCACCTCACAATGAT 544
 QY 182 GluAlaValAlaPhePheCysAsnLeuMetAspMetValGluHisTrpSerIlyAspIlySGly 201
 |||
 Db 545 GAGGCCCTGTCTCTCAACCTCATGACATGATGAGATGATACAGCAAGCAAGAGGC 604
 QY 202 AlaIleCysThrIlySerLeuValArGValProIlyArGValIlySerIlyThrIlySerAlaGlu 221
 |||
 Db 605 GCTATCTGCACCAAGCTGTGAGACCAAGCGGAACAGCGGACCAAGTGGCGGAGAG 664
 QY 222 GluLeuAlaArGValGlyTrpLeuLeuAsnLeuGlnHisIleuThrIleGlyAlaGlnIle 241
 |||
 Db 665 GAGCTGGCCAGGGGGCTGTACTGAACCTGCAGATTTGACATTTGGGAGCACAGATC 724
 QY 242 GlyGluGlyIlyPheGlyAlaValLeuGlnGlyIlyTrpLeuGlyIlyValAlaVal 261
 |||
 Db 725 GGAGAGGAGAGATTGGAGCTGCTGCAAGGAGTGAAGTCTGGGGCAAAAG---TGGCTG 781
 QY 262 LysAsnIleLysCysAspValIleThrAlaGlnAlaPheLeuAspGluTrpAlaValMetThr 281
 |||
 Db 782 AAGAAATATCAAGTGTGATGTGACAGCCAGGCTTCTCGAGCAGAGAGCGCTCATGACG 841
 QY 282 LysMetGlnHisGluAsnLeuValArGLeuLeuGlyValIleLeuHisGlnGlyLeuTyr 301
 |||
 Db 842 AAGATGAAACACAGAACTGTGCTGCTCTGCGGTGATCTCCAGCCAGGGGCTGTAC 901
 QY 302 IleValMetGluHisValSerIlyGlyAsnLeuValAsnPheLeuArGTrpArGlyArg 321
 |||
 Db 902 ATTGTATGAGAGCAGCAGTGAAGGCAAGGCAACTGTGTAATCTTCTGCGAGCCCGGGTGA 961
 QY 322 AlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMetGlu 341
 |||
 Db 962 GCCCTGTGAACCCGCTGAGCTGCTGCACTTTCTCTGCACTGGCGAGGGGCAATGAG 1021
 QY 342 TyrLeuGluSerIlySerIlyLeuValHisArGAspLeuAlaAlaArgAsnIleLeuValSer 361
 |||
 Db 1022 TACTTGGAGAGCAAGAACTTTGTGCACCGGACCTAGCCCGCCCAATCTCTGCTCA 1081
 QY 362 GluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArGlySGlyLeu 381
 |||
 Db 1082 GAGAGAGTGTGGCCAAAGTCAAGCTTGGCTGGCCAAAGCCGAGGGGAGGCTA 1141
 QY 382 AspSerSerArgLeuProValIlyTrpThrAlaProGluAlaLeuLysHisGlyLysPhe 401

Db 1142 GACTCAAGCCGGCTGCCCTCAAGTGACGCGCCCGCAGCTCTCAAAACCGGAAGTTC 1201
 QY 402 ThrSerIlySerAspValIleTrpSerPheGlyValLeuLeuTrpGluValPheSerTyrGly 421
 |||
 Db 1202 ACCAGCAAGTCGATGTCTGTGAGTTTGGGGTCTGCTGTGGAGGCTTTCTCATATGA 1261
 QY 422 ArgAlaProIlyProIlyMetSerLeuIlySGlyValSerGluAlaValGluIlySGlyTyr 441
 |||
 Db 1262 CGGGCTCCGTAACCTAAATGTCACTGAAGAGGTGTGGAGGCGGTGGAGAGCGGTAC 1321
 QY 442 ArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCysTrp 461
 |||
 Db 1322 CGCATGAAACCCCGAGGGGCTGTCCAGGCGCCGTGACGCTCTCATAGACCACTGCTG 1381
 QY 462 GluAlaGluProAlaArGArGProProPheArGlySerAlaGluIlySerLeuAlaArgGlu 481
 |||
 Db 1382 GAGGCAAGCCCGCCCGCCGACACCTTCCGCAACTGGCCGAAAGCTGCGCCGGAG 1441
 QY 482 LeuArGSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThrSer 501
 |||
 Db 1442 CTACGCAAGTCAGAGTGGCCCGACCTCGCTCAGGGCAGAGCCGCGAGCTCCACTCG 1501
 QY 502 ProArgSerGlnGluPro 507
 |||
 Db 1502 CCCGGAAGCCAGAGGCC 1519
 RESULT 11
 AR084686
 LOCUS AR084686 1987 bp DNA linear PAT 01-SEP-2000
 DEFINITION Sequence 1 from patent US 5981201.
 ACCESSION AR084686
 VERSION AR084686.1 GI:10011456
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1987)
 AUTHORS Avraham, H. and Groopman, J.E.
 TITLE Methods of detection and treatment of breast cancer
 JOURNAL Patent: US 5981201-A 1 09-NOV-1999;
 FEATURES
 source location/Qualifiers
 1..1987
 BASE COUNT 369 a 628 c 672 g 318 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,89e-153 Length: 1987
 Score: 2591.50 Matches: 502
 Percent Similarity: 99.21% Conservative: 1
 Best Local Similarity: 99.01% Mismatches: 3
 Query Match: 97.02% Indels: 3
 DB: 6 Gaps: 1
 US-09-977-260-2 (1-507) x AR084686 (1-1987)
 QY 1 MetAlaGlyArGlySerLeuValSerTrpArGAlaPheHisGlyCysAspSerAlaGlu 20
 |||
 Db 263 ATGGCGGGGAGAGGCTCTGTGTTCTCTGGCGGGCATTTACAGCGCTGTATTTCTGTAG 322
 QY 21 GluLeuProArGValSerProArGpHleuArGAlaTrpHisProProProValSerAla 40
 |||
 Db 323 GAATCTCCCGGGGTGAGCCCGCTTCTCTCGAGCTGGGACCCCGCTCCGCTCAGCC 382
 QY 41 ArgMetProthArGArGTrpAlaProGlyThrGlnCysIleThrIlyScySGluHisThr 60
 |||
 Db 383 AGGATGCCAACGAGGGGCTGGGCGCGGACCACTGTATCAACCAATGTGAGCACACC 442
 QY 61 ArgProIlySerProGlyGluLeuAlaPheArGlySGlyAspValAlaThrIleLeuGluAla 80
 |||
 Db 443 CGCCCAAGCCAGGAGAGCTGGCTTCCGCAAGGGCAGCTGTCTCAACATCTCTGGAGGCC 502

QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGlyLeu 100
 |||||
 Db 503 TGGAGAAACAAGAGCTGTACCGGCTGACAGACCAAGTGGACAGAGGGCTCTG 562
 QY 101 AAlaIaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 |||||
 Db 563 GCAGCTGGGGCGCTGGCGGACGGGGCCCTCTCCACAGACCCCAAGCTCAGCTCATG 622
 QY 121 ProTrpPheHisGlyLysHisSerGlyGlnGluAlaGlnGlnLeuGlnProProGlu 140
 |||||
 Db 623 CCGTGGTCCACGGGAGATCTGGGCGAGAGGCTGTCCAGCAGCTGCGAGCCCTCCGAG 682
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysVal 160
 |||||
 Db 683 GATGGGCTGTCTCGTGGGAGTCCGGCGGCCACCCCGGAGCTACGCTCTGTGGTG 742
 QY 161 SerPheGlyArgAspValHisHisTyrArgValLeuHisArgAspGlyHisLeuThrTle 180
 |||||
 Db 743 AGCTTTGGCGCGCGCTCATCCATCCGCGTGTCTGACCGCGACGCGCCACCTCACATC 802
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys 200
 |||||
 Db 803 GATGAGCCCTGTCTTCTGTCAACCTCATGAGACTGTGGACATTACAGCAAGAGAG 862
 QY 201 GAlaAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 |||||
 Db 863 GGGCGCTATCGACCAAGCTGGTGTAGACCAAGCGGAACAGCGGACCAAGTCCGCCGAG 922
 QY 221 GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 |||||
 Db 923 GAGAGCTGGCCAGGGCGGCTGTACTGACCACTGACGATTTGACATTGGGAGCAG 982
 QY 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyGlyTyrLeuGlyGlnLysValAla 260
 |||||
 Db 983 ATGGAGAGGAGAGAGTTGGAGCTGTCTGTGAGGTGATGACCTGGGCGCAAAAGTGGCC 1042
 QY 261 ValLysAsnLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
 |||||
 Db 1043 GTGAGAAATATCAAGTGTGATGTGACAGCCAGGCTTCCGTGAGAGAGCGCGTCA 1102
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
 |||||
 Db 1103 ACGAAGATGCACGAGAACCTGGTGTCTGTCTGGCGGTGATCCGACCGAGGGCTG 1162
 QY 301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 |||||
 Db 1163 TACATTTCATGGAGCGCTGTGACAGGCAACCTGTGTGACTTTCGCGGACCCGGGCT 1222
 QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet 340
 |||||
 Db 1223 CGAGCCCTCGTGAACACCGCTCAGCTCTGCAAGTTTCTGTGACGCGGCGAGGGCATG 1282
 QY 341 GluTyrLeuGluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 |||||
 Db 1283 GAGTACCTGAGACCAAGAGCTGTGTGACCGGCACTGGCGCGCCCAACATCTCTGTC 1342
 QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
 |||||
 Db 1343 TCGAGAGACCTGTGGCGCAAGGTCAGGACTTTGGCTGTGGCCAAAGCCGAGCGGAAGGG 1402
 QY 381 LeuAspSerSerArgLeuProValLysTyrThrAlaProGluAlaLeuLysHisGlyLys 400
 |||||
 Db 1403 CTGAGACTCAAGCGCGCTGCCGTCAGTGGAGCGCGCGGCTCTCAAAACAGGG - 1459
 QY 401 PheThrSerLysSerAspValTyrPheSerPheGlyValLeuLeuTyrGluValPheSerTyr 420
 |||||
 Db 1460 TTACACGCAAGTGTGATGTCTGAGATTGGGGTGTCTGTGGGAGGCTTCTCTCATAT 1519
 QY 421 GluArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 |||||
 Db 1520 GCGAGGGCTCGTACCTAAATATCTCACTGAAGAGGTGTGCGAGGCGCGTGGAGAGAGGG 1579
 QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460

Db 1580 TACCGATGGAACCCCGGAGGGCTGTCCAGGGCCCGGTGACGCTCTCATAGAGAGCTGC 1639
 QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
 |||||
 Db 1640 TGGGAGGACAGAGCGGCC - CGCCGCGCACCTTCGCGAAACTGGCGGAGAGCTGGCCGG 1698
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 |||||
 Db 1699 GAGCTTACGACATGTCAGGTGTCCCGACCTCGTCTCAGAGGAGAGCGCGAGGTTC - ACC 1757
 QY 501 SerProArgSerGlnGluPro 507
 |||||
 Db 1758 TCGCCCGGAGGACGAGAGCC 1778
 RESULT 12
 AR204704 1987 bp DNA linear PAT 20-JUN-2002
 LOCUS AR204704
 DEFINITION Sequence 1 from patent US 6368796.
 ACCESSION AR204704
 VERSION AR204704.1 GI:21502099
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCES
 1. (bases 1 to 1987)
 AUTHORS Avraham, H. and Groopman, J. E.
 TITLE Methods of detection and treatment of breast cancer
 JOURNAL Patent: US 6368796-A 1 09-Apr-2002;
 FEATURES
 source location/Qualifiers
 1. 1987
 BASE COUNT 369 a 628 c 672 g 318 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.89e-153 Length: 1987
 Score: 2591.50 Matches: 502
 Percent Similarity: 99.21% Conservative: 1
 Best Local Similarity: 99.01% Mismatches: 3
 Query Match: 97.02% Indels: 1
 DB: 6 Gaps: 1
 US-09-977-260-2 (1-507) x AR204704 (1-1987)
 QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 |||||
 Db 263 ATGGCGGGGAGAGCTCTGTGTTCTGTGGCGGCAATTTCACGCTGTGATTCGTGAG 322
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisPropProValSerAla 40
 |||||
 Db 323 GAACCTCCCGGGGTGAGCCCGCTCTCTCGAGCTGGACACCCCTCCGCTCTCAGCC 382
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
 |||||
 Db 383 AGATGCGCAAGAGAGCGCTGGCGCCCGGACACCAAGTATACCAAAATGAGACACACC 442
 QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValThrIleLeuGluAla 80
 |||||
 Db 443 CGCCCAAGCGAGGGAGGTGGCTTCGGAAGGCGACGTCGACCATCTGGAGGCC 502
 QY 81 CysGluAsnLysSerTrpTyrArgValLysHisHisThrSerGlyGlnGlyLeuLeu 100
 |||||
 Db 503 TGGGAGAACAGAGCTGTACCGCTGTACAGCACACACCAAGTGGAGACGGAGGGCTCTG 562
 QY 101 AAlaIaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 |||||
 Db 563 GCAGCTGGGCGCTGGCGGAGGGAGGCCCTCTCCGACAGACCCCAAGCTCAGGCTCATG 622
 QY 121 ProTrpPheHisGlyLysHisSerGlyGlnGluAlaGlnGlnLeuGlnProProGlu 140
 |||||
 Db 623 CCGTGGTCCACGGGAGATCTCGGCGAGAGGCTGTCCAGAGAGCTGACGAGCTCCCGAG 682

141 AspglyLeuPheLeuValArgGlySerAlaArgHisProGlyAspTyrValLeucysVal 160
|||||
683 GATGGCTGTTCCTGGTGGGAGTCCGGCCACCCCGGAGACACTCTCTGTCCTG 742
QY SerPheGlyArgAspValIleHisTyrArgValIleuHisArgAspGlyHisLeuThrIle 180
|||||
743 ACCTTTGGCCGGGAGCTCATCCACTACCCGCTGCTACCCGGGAGCGGCACTTCACAAATC 802
Db
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerIlysAspLys 200
|||||
803 GATGAGGCGCGTGTCTTCTGCAACCTCATGAGCATGTGGAGCATATACAGCAAGCAAG 862
Db
QY 201 GlyAlaIleCysThrIlysLeuValArgProLysArgLysHisGlyThrIysSerAlaGlu 220
|||||
863 GGGCTATCTGCACCAAGCTGTGAGACCAAAAGCGGAAACAGGGGCCCAAGTCCGGCCGAG 922
Db
QY 221 GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGluHisIleuThrLeuGlyAlaGln 240
|||||
923 GAGGAGCTGGCCAGGGGGGCTGCTTACTGAACTGACACATTTGACATTTGGGAGCAG 982
Db
QY 241 IleGlyGluGlyLeuPheGlyAlaValLeuGlnGlyLysIleuGlyLysValAla 260
|||||
983 ATCGGAGAGGAGAGATTGGAGCTGTCTGCAAGGAGTACTGAGGCAAAAGTGGCC 1042
Db
QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
|||||
1043 GTGAAGAATATCAAGTGTATGTACAGCCAGGCTTCTGACAGACAGCGCCGCTCAG 1102
Db
QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleuHisGlnGlyLeu 300
|||||
1103 AGGAGATGCACACAGAGAACTGTGCTGCTCTGCTGGGTATCTGACACAGGGGCTG 1162
Db
QY 301 TyrIleValMetGluHisValSerIysGlyAsnLeuValAsnPheLeuArgThrArgLys 320
|||||
1163 TACATTTGTATGAGAGACGTGAGCAAGGCAACTGTGAACTTCTGCGGAGCCCGGCT 1222
Db
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluGlyMet 340
|||||
1223 CGAGCCCTGTGAAACACCGCTGACCTGCGAGTTTCTCTGACGTGGCCAGGGCAG 1282
Db
QY 341 GluTyrLeuGlySerIysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
|||||
1283 GAGTACCTGGAGACAGAAAGCTTGTGCAACCCGACCTTGGCCCGGCAACATCTCTGCTC 1342
Db
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
|||||
1343 TCAGAGGACCTGCTGGCCAAAGTCAAGCAGCTTGGCTGGCCAAAGCCGAGGGAAGGG 1402
Db
QY 381 LeuAspSerSerArgLeuProValLysTyrThrAlaProGluAlaLeuLysHisGlyLys 400
|||||
1403 CTAGACTCAAGCCGCTGCCCTCAAGTGGAGCGGCCCGGAGGCTCTCAAAACAGGG -- 1459
Db
QY 401 PheThrSerLysSerAspValTyrSerPheGlyValIleLeuThrGluValPheSerTyr 420
|||||
1460 TTCACACAGCAATCGGATGCTGGAGTTTGGGGTGCTCTCTGGGAGGTCTTCTCATAT 1519
Db
QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
|||||
1520 GGACGGGCTCCTACCTCAATATGTCAGAAAGAGTGTGAGAGCCGTGAGAAAGGG 1579
Db
QY 441 TyrArgMetGluProProGluGlyLysProGlyProValHisValLeuMetSerCys 460
|||||
1580 TACCGCATGGAACCCCGGAGGGCTGTCCAGGCGCCCGTGCACAGTCTCTCATGAGCTGC 1639
Db
QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
|||||
1640 TGGGAGCGAGACCCGCC -CGCGCGCACCTTCCGCAAACTGGCCGAGAGAGCTGGCCCG 1698
Db
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGluAspAlaAspLysThr 500
|||||
1699 GAGCTACGAGAGTGCAGGTGCCCAAGCTCTGCTCAGGCGAGGAGCGCCAGCGTCC -ACC 1757
Db
QY 501 SerProArgSerGlnGluPro 507

Db 1758 TCGCCCGAAGCCAGAGGCC 1778
|||||
RESULT 13
HUMANK
LOCUS HUMANK 1987 bp mRNA linear PRI 11-MAR-1994
DEFINITION Human tyrosine kinase (MATK) mRNA, complete cds.
ACCESSION L18974
VERSION L18974.1 GI:455449
KEYWORDS cytoplasmic protein; tyrosine kinase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1987)
Bennett,B.D., Cowley,S., Jiang,S., London,R., Deng,B., Grabarek,J.,
Groopman,J.E., Goeddel,D.V. and Avraham,H.
AUTHORS Identification and characterization of a novel tyrosine kinase from
megakaryocytes
JOURNAL J. Biol. Chem. 269 (2), 1068-1074 (1994)
MEDLINE 94117408
PubMed 8288563
COMMENT Original source text: Homo sapiens cDNA to mRNA.
FEATURES
source
Location/Qualifiers
1..1987
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="CMK11-5"
/cell_type="megakaryocyte"
/dev_stage="adult"
/tissue_1lb="lambda gt11"
1..1987
/gene="MATK"
<1..63
/gene="MATK"
263..1846
/gene="MATK"
/codon_start=1
/product="tyrosine kinase"
/protein_id="AA16703.1"
/db_xref="GI:459550"
/translation="MAGKSLYSMAFHCDSAEELPRVSPFLFAMHPPEVSAMPT
RMAPGTCTICENTRPKPGELARGLDVVITLDECEKSWRVKHNHSGEGLLAA
GALRGEALSDAPKLSLMPWFGKISGDEAVOQLPPEDEGLFLVESARHPDDVLCV
SGRQVYIHRYLRDGHLLTIDAVFECNLMADVHSKDKAICTKLVKPKRHGTS
AEEELARGLMLNOLHITLGAOIGGEGCNLYOGKLVKNIKCDVNRQAFDE
TAVMTKMOHENTVRLGLVITLHGLTYMEHVSKGLVNFRTGRGALVNTQLDPSL
HVAEHEYLIESKLLVHRDLAARNILYSEDLVAKVSDFLAKERKGLDSRLPVKMTA
PEALKHGFSTKSDVNSFVLLMEVFSYGAPYPKSLKEVSEAVEKGYRMEPEGCPG
PVHVLMSQMEAEPPAGHPSAMWPSRWSYAVOVPPPSOGRTPTVHLAPKPAITP
FGPWPORTEREVSAAWG"
568
polyA_site
BASE COUNT 369 a 628 c 672 g 318 t
ORIGIN
Alignment Scores:
Pred. No.: 1,89e-153 Length: 1987
Score: 2591.50 Matches: 502
Percent Similarity: 99.21% Conservative: 1
Best Local Similarity: 99.01% Mismatches: 3
Query Match: 97.02% Indels: 3
DB: 9 Gaps: 1
US-09-977-260-2 (1-507) x HUMANK (1-1987)
QY 1 MetAlaGlyArgGlySerLeuValSerTyrPheAlaPheHisGlyCysAspSerAlaGlu 20
|||||
Db 263 ATCGCGGGCGAGGCTCTGCTGCTGCGGCGCATTTCAAGCGCTGATCTGCTGAG 322
|||||
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTyrPheProProValSerAla 40
|||||

Db 323 GAACCTCCCGGGTGAGCCCCCGCTTCGAGACCTGGACACCCCGCTCCGCTTCAGCC 382
 QY ArgMetProThrArgTgTTPalaProGlyThrgInGlySileThryCysGluHisThr 60
 Db 383 AGATGCGCAACGAGCGCTGGCGCCCGGCGACCCAGTGTATACCAATGCGACACACC 442
 QY ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
 Db 443 CGCCCCAAGCGAGGAGCTGGCGCTTCGCAAGAGGCGACGTGTGCACATCCTGGAGGCC 502
 QY CysGluAsnLysSerTTPyTArgValLysHisThrSerGlyGlnGlyLeuLeu 100
 Db 503 TCGCAGACACAGACTGTACCGCGTCAACACACACACAGTGTACAGAGGGGCTCTGTG 562
 QY AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 Db 563 GCAGCTGGGCGCGTGGCGGAGGAGGAGGCGCTTCGCGACAGCCCAAGCTCAGCTTATG 622
 QY ProTPrheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProGlu 140
 Db 623 CCGGTGTCCACGGGAGATCTCGGGCGAGAGGCTGTCCAGCAGCTGCAGCTCCGCGAG 682
 QY AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal 160
 Db 683 GATGGCGTGTCTCGTGTGCGGAGTCCGCGCGCACCCCGCGACTACGCTGTGTGCGTG 742
 QY SerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
 Db 743 AGCTTTGGCGCGACGTCATCCACTACCGCGTGTGTGACGCGCGGCGCACCTCACAATC 802
 QY AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys 200
 Db 803 GATGAGCGCGGTCTCTCTGCAACCTCATGATGATGTGAGCATATACAGCAAGACAG 862
 QY GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 Db 863 GGGCGTATCTGTCCACCAAGCTGTGTGACCAAGGGAACACGCGGACCAACTCGCGCGAG 922
 QY GluGluLeuAlaArgAlaGlyTTPLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 Db 923 GAGGAGCTGGCGAGGGGCGGTGTGTACTGAACTGCAGCATTTGACATTTGGAGCACAG 982
 QY IleGlyGlnGlyLupheGlyAlaValLeuGlnGlyLupTyrLeuGlyGlnLysValAla 260
 Db 983 ATCGAAGAGGAGGAGTGTGAGCTGTCTCGAGGGTGAATCACTGGGCGCAAAAGGTGACC 1042
 QY ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
 Db 1043 GTGAAGAATATCAAGTGTGATGTACAGCCCGGCTCTCGAGCAGAGCGCGGTGATG 1102
 QY ThrLysMetGlnHisGluAsnLeuValArgLeuGlnGlyAlaIleLeuHisGlnGlyLeu 300
 Db 1103 ACGAAGATCGAACACGAGAACCTGTGCTGTGGCGGTGATCTTCACCAAGGGGCTG 1162
 QY TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 Db 1163 TACATTTGATGAGCAGCAGTGTGACAGGCAACTGTGTGACTTTCGCGGACCGGGGT 1222
 QY ArgAlaLeuValAsnThrAlaGlnLeuGlnInPheSerLeuHisValAlaGlnGlyMet 340
 Db 1223 CGAGCCCTGTGAACCGCGTCAAGCTCTGTGAGTTTTCTGTGACGTGGCGGAGGGATG 1282
 QY GluTyrLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 Db 1283 GAGTACTCGAGACCAAGACTGTGTACCGGCACTGTGGCGCGCCCAACATCTCTGTGTC 1342
 QY SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
 Db 1343 TCAAGAGACTGTGTGGCCAAAGGTACGAGACTTTGGCTGGCGCAAAAGCGGAGAGGGG 1402
 QY LeuAspSerSerArgLeuProValLysTyrThrAlaProGluAlaLeuLysHisGlyLys 400
 Db 1403 CTGAGCTCAAGCCGCGTGGCGGTCAAGTGTGAGCGGCGCGAGGCTCTCAAAACAGCGG-- 1459

QY 401 PheThrSerLysSerAspValTTPSerPheGlyValLeuLeuTrrpGluValPheSerTyr 420
 Db 1460 TTCACCGCAAGTGTGATGTGTGAGTGTGGGGTGTCTGCTGTGGAGGCTTCTCATAT 1519
 QY 421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 Db 1520 GAGCGGGCTCCGTACCTAAATGTCACTGAAAGAGGTGTGTGAGCGCGGTGGAGAGGGG 1579
 QY 441 TyrArgMetGluProProGlnGlyCysProGlyProValHisValLeuMetSerSerCys 460
 Db 1580 TACCGCATGGAAACCCCGGAGGGCTGTTCAGGCGCCCTGTGACGCTCATGACGAGCTGC 1639
 QY 461 TrpGluAlaGluProAlaArgArgProProPheArgGlySerLeuAlaGluLysLeuAlaArg 480
 Db 1640 TGGGAGCAGAGCCGCC-CCGCGCGCACCTTCGCGCAAACTGCGCGAGAAAGCTGGCCCGG 1698
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 Db 1699 GAGCTACGACAGTGTGAGGTGCGCCAGCTCGGTCTCAGGCGAGAGCGGAGGCTCC-ACC 1757
 QY 501 SerProArgSerGlnGluPro 507
 Db 1758 TCGCCCGCAAGCGAGAGCC 1778
 RESULT 14
 HSM802388
 LOCUS HSM802388 1804 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA; cDNA DKFZp434N1212 (from clone DKFZp434N1212); partial cds.
 ACCESSION AL137754
 VERSION AL137754.1 GI:6808456
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1804)
 Duesterhoeft,A., Lauber,J., Mewes,H.W., Weill,B. and Wiemann,S.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUN-2000) MTPS, Am Kiofepersplitz 18a, D-82152 Martinsried, GERMANY
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by Olegen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp434N1212) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de. Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
 FEATURES
 source
 1..1804
 location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="19p13"
 /clone="DKFZp434N1212"
 /tissue_type="testis"
 /clone_id="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"
 /dev_stage="adult"
 1..1804
 /gene="DKFZp434N1212"
 <1..1574
 /gene="DKFZp434N1212"
 /note="protein-tyrosine kinase (EC 2.7.1.112) (Homo sapiens)"
 /product="hypothetical protein"
 /codon_start=3
 /protein_id="CAB70906.1"
 /db_xref="GI:6808457"
 /translation="GNGATRSKLCSDRLGVCSSRLAEALIGAGAGRGSGAPPEQVSPR

Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2073)
 Isogai, T., Otsuki, T. and Sugiyama, T.
 Direct Submission
 Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Katsurazu, Chiba 293-0812, Japan
 (E-mail: genomics@helix.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.

COMMENT

FEATURES
 source
 1..2073
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="FEBRA202027"
 /tissue_type="brain"
 /clone_lib="FEBRA2"
 /dev_stage="fetus"
 /note="Cloning vector: pME18SFL3"

BASE COUNT 415 a 634 c 703 g 321 t
 ORIGIN

Alignment Scores:

Pred. No.: 2,696-150 Length: 2073
 Score: 2541.50 Matches: 489
 Percent Similarity: 94.06% Conservative: 2
 Best Local Similarity: 93.68% Mismatches: 6
 Query Match: 95.15% Indels: 25
 DB: 9 Gaps: 2

US-09-977-260-2 (1-507) x AK055395 (1-2073)

QY 11 ArgAlaPheHisGlyCys----- 16
 DB 297 AAAGCTTCATGCGGGTCTAGAGCATCTTAATGTCCTTCATGACGAGACATTTCCCT 356
 QY 17 ---AspSerAlaGluGluLeuProArg-----Val 25
 DB 357 GCTGAACGAGAGGAGGACAGCTTCGGGAGGGACCCGTGTCAGCAGCAGCTTCTGTG 416
 QY 26 SerProArgPheLeuArgAlaTrpHisProProProValSerAlaArgMetProThrArg 45
 DB 417 AGCCCCCGCTTCCTCCAGCCGCGACCCCTCCCGTCTCAGCCAGGATGCCAACGAGG 476
 QY 46 ArgTrpAlaProGlyThrGlnCysIleThrIlyCysGluHisThrArgProLysProGly 65
 DB 477 CGCTGGGCGCCGCGACCCAGTGTATCACCAATGTCGAGCACACCCGCCCAAGCCAGGG 536
 QY 66 GluLeuAlaPheArgGlyGlyAspValAlaThrIleLeuGluAlaCysGluAsnLysSer 85
 DB 537 GAGCGTGGCTTCGGAAGGGCGACGTGTGACCATCTGAGGCGCTGCGAGAACAGAGC 596
 QY 86 TrpTrpArgValIlyHisIleThrSerGlyGlnGluGlyLeuLeuAlaAlaGlyAlaLeu 105
 DB 597 TGGTACCGGCTCAACACACACACAGTGAGACGAGGGGCTGCTGGCAGCTGGGGCGTGC 656
 QY 106 ArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetProTrpPheHisGly 125
 DB 657 CGGAGCGGGAGGCGCTTCGCCAGACCCCAAGCTCAGCTCATGCCGTGGTTCACAGGG 716
 QY 126 LysIleSerGlyGlnGluAlaValGlnGlnIleuGlnProProGluAspGlyLeuPheLeu 145
 DB 717 AAGATCTGGGCGGAGGCTGTCCAGACGCTGACGCTCCGAGAGATGGGCTGTTCG 776

QY 146 ValArgLysSerAlaArgHisProGlyAspTrpValLeuCysValSerPheGlyArgAsp 165
 DB 777 GTGGGGAGATCCGGCGCCACCCCGGAGCTAGCTCTGTGGTGAGACTTTGGCGCGGAC 836
 QY 166 ValIleHisTrpArgValLeuHisArgAspGlyHisLeuThrIleAspGluAlaValPhe 185
 DB 837 GTCATCCACATACCCCGCTGACACCCGAGGGCCACTCCAAATCGATGAGCGCTGTTCC 896
 QY 186 PheCysAsnLeuMetAlaPmetValGluHisIleThrSerLysAspLysGlyAlaIleCysThr 205
 DB 897 TTCTGCAACCTCATGACATGATGTGACCATTTACGACAAAGGACAGGGCGCTATCTGCACC 956
 QY 206 LysLeuValArgProLysArgLysHisGlyThrLysSerAlaGluGlnGluLeuAlaArg 225
 DB 957 AAGCTGTGAGACCAAAAGCGGAAACAGGGGACCAAGTCGGCGGAGGAGCTGGCCAGG 1016
 QY 226 AlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGluGlyLeu 245
 DB 1017 GCGGGCTGGTATTGAACCTGCACCATTTGTGACATTTGGAGCCACGATCGGAGGAGAG 1076
 QY 246 PheGlyAlaValLeuGlnGlyLeuTrpLeuGlyGlnLysValAlaValLysAsnIleLys 265
 DB 1077 TTTCGAGCTCTCTCCAGAGGTGAGTACTCTGGGCGCAAAAGTGGCCGTGMAAATATCAAG 1136
 QY 266 CysAspValThrAlaGlnAlaPheLeuAspGlyThrAlaValMetThrLysMetGlnHis 285
 DB 1137 TGTGATGTGACAGCCAGGCTTCCTGAGACAGAGCGCCGTCATGACGAATGCAACAC 1196
 QY 286 GluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeuTrpIleValMetGlu 305
 DB 1197 GAGAACCTGGTCTCTCTGGGCGTATCTTGACCAAGGGGCTGTATTTGTATGTAGAG 1256
 QY 306 HisValSerLysGlyAsnLeuValAsnPheLeuArgTrpArgLysArgAlaLeuValAsn 325
 DB 1257 CACGTGACCAAGGGCAACCTGTGATCTTGTGGAGACCGGGGTGACCGCTGTGAGAC 1316
 QY 326 ThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMetGlyTrpLeuGlnSer 345
 DB 1317 ACCGCTCAGCTCTGCAATTTCTCTGACGCTGCGCAGGCGCATGAGTACTGTGAGAGC 1376
 QY 346 LysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuVal 365
 DB 1377 AAGAGCTTGTGCACCGGACCTGTGCGCCGCAACATCTGTCTCAGAGACCTGTGTG 1436
 QY 366 AlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArg 385
 DB 1437 GCCAAGTCCAGCAGCTTTGGCTGCGCAAAACCCAGCGGAGGGCTGACTCAAGCCGG 1496
 QY 386 LeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLysPheThrSerLysSer 405
 DB 1497 CTGCCGTCAGATGGACGCGCCGCGAGGCTCTCAAAACGCGGAAGTTCCACGACAGATCG 1556
 QY 406 AspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrpGlyArgAlaProTrp 425
 DB 1557 GATGCTGAGATTGGGGTCTGCTGTGGAGGCTTCTCAATATGAGCGGGCTCCGATC 1616
 QY 426 ProLysMetSerLeuLysGluValSerGluAlaValGluLysGlyTrpArgMetGluPro 445
 DB 1617 CCTAAATGTCACTAAAGAGGTGTGAGAGCCGCTGGAAGAGGGGTACCGCATGGAACCC 1676
 QY 446 ProGluGlyCysProGlyProValHisValLeuMetSerCysTrpGluAlaGluPro 465
 DB 1677 CCCGAGGGCTGTCCAGGCGCCGCTGCACTCTCATAGACAGCTGTGAGAGCAGAGGCC 1736
 QY 466 AlaArgArgProProPheArgGlySerLeuAlaGluLysLeuAlaArgGluLeuArgSerAla 485
 DB 1737 GCCCGCCGACACCTTCCGAACTGGCGGAGAGCTGGCCCGGAGACTACGACGTGCA 1796
 QY 486 GlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThrSerProArgSerGln 505
 DB 1797 GGTGCCCAAGCTCGTCTCAGGAGGACGCGGACGCGCTCACCTGCCCCGGAAGCCAG 1856
 QY 506 GluPro 507

Mon Aug 4 14:40:04 2003

us-09-977-260-2.rge

Page 18

|||||
Db 1857 GAGCC 1862

Search completed: August 1, 2003, 21:43:10
Job time : 6062 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:35:28 ; Search time 9490 Seconds

(without alignments)
8621.628 Million cell updates/sec

Title: US-09-977-260-1

Perfect score: 2000

Sequence: 1 ctgcctccaagtgtgcagc.....attctaagactctaaanaa 2000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenBank:

1: gb.ba:**

2: gb.htg:**

3: gb.in:**

4: gb.om:**

5: gb.ov:**

6: gb.pat:**

7: gb.ph:**

8: gb.pl:**

9: gb.pr:**

10: gb.ro:**

11: gb.sts:**

12: gb.sy:**

13: gb.un:**

14: gb.vl:**

15: em.ba:**

16: em.fun:**

17: em.hum:**

18: em.in:**

19: em.mu:**

20: em.om:**

21: em.or:**

22: em.ov:**

23: em.pat:**

24: em.ph:**

25: em.pl:**

26: em.ro:**

27: em.sts:**

28: em.un:**

29: em.vl:**

30: em.htg.hum:**

31: em.htg.in:**

32: em.htg.other:**

33: em.htg.mus:**

34: em.htg.pln:**

35: em.htg.rod:**

36: em.htg.mam:**

37: em.htg.vrt:**

38: em.sy:**

39: em.htgo.hum:**

40: em.htgo.mus:**

41: em.htgo.other:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2000	100.0	2000	6	AR179662
2	2000	100.0	2000	6	AR274487
3	1971.2	98.6	2180	6	BC000114
4	1971.2	98.6	2180	6	BC000114 Homo sapi
5	1963.4	98.2	1997	6	BC003109
6	1909.8	95.5	1987	6	BD16710
7	1909.8	95.5	1987	6	BD16710 Method of
8	1909.8	95.5	1987	6	AR084686
9	1909.8	95.5	1987	6	AR204704
10	1909.4	95.5	1968	9	HDMATK
11	1903.4	95.2	1942	6	HSYLYNK
12	1669.2	83.5	1804	9	AR053292
13	1663	83.2	2073	9	AR053292
14	1651.4	82.6	1744	9	AR053292
15	1519.4	76.0	1521	6	AR053291
16	1398	69.9	1398	6	AR053290
17	1377	68.8	1713	6	AX644057
18	1103.2	55.2	1734	10	MUSNRK
19	1071.4	53.6	1911	10	MUSCTR
20	1068.6	53.4	1651	10	MMU05210
21	1066	53.3	1656	10	S77473
22	1060	53.0	1838	10	RATBATK
23	738	36.9	738	6	AR053289
24	547.6	27.4	2187	6	AX330746
25	547.6	27.4	2187	9	HSCSRCKIN
26	547.6	27.4	2420	9	HSCYLCYK
27	537	26.9	1682	9	AY007162
28	521	26.1	1353	5	CHRSCKCA
29	506	25.3	2125	10	REYTKI
30	496.4	24.8	2292	10	BC018394
31	496.4	24.8	2363	10	BC052006
32	491.6	24.6	1779	10	MMU05247
33	455	22.8	16389	9	AX644059
34	455	22.8	43190	6	AC005777
35	440	22.0	457	9	H00687453
36	372.2	18.6	1806	5	AF052430
37	281.6	14.1	2205	3	AK113892
38	272.4	13.6	1733	3	AB006558
39	236.2	11.8	245	9	H006874S04
40	225	11.2	225	6	AR053288
41	220	11.0	233	9	H006874S05
42	217.4	10.9	233	9	H006874S02
43	216.4	10.8	1589	9	HS007236
44	215.2	10.8	1982	5	GGYRKA
45	213.6	10.7	3299	6	BD160548

ALIGNMENTS

RESULT 1	AR179662	2000 bp	DNA	1linear	PAT 20-APR-2002
LOCUS	AR179662				
DEFINITION	Sequence 1 from patent US 6326469.				
ACCESSION	AR179662.1				
VERSION	AR179662.1				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 2000)				
AUTHORS	Ullrich,A., Gishizky,M. and Sures,I.				
TITLE	Megakaryocytic protein tyrosine kinases				
JOURNAL	Patent: US 6326469-A 1 04-DEC-2001;				
FEATURES	Location/Qualifiers				

source 1..2000
/organism="unknown"
BASE COUNT 375 a 632 c 673 g 320 t
ORIGIN
Query Match 100.0%; Score 2000; DB 6; Length 2000;
Best Local Similarity 100.0%; Pred. No. 3.2e-289;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCGCTCCAAAGTTGTGACGCGGACCGCTCGGGGTGTGACGCGGCTCGGGAGGCC 60
DB 1 CTGCTCCAAAGTTGTGACGCGGACCGCTCGGGGTGTGACGCGGCTCGGGAGGCC 60
QY 61 TCCTGG 120
DB 61 TCCTGG 120
QY 121 AGGCTGGGTCCAGTGGGACCGACGCTCCCTACCTCTGTGACGCGGCTGGGGTGGCA 180
DB 121 AGGCTGGGTCCAGTGGGACCGACGCTCCCTACCTCTGTGACGCGGCTGGGGTGGCA 180
QY 181 GGGCATTCCAGCGTCCCGACGCTGTGACCACTTGTCTGCTGCTCTACCTGCTCTGAG 240
DB 181 GGGCATTCCAGCGTCCCGACGCTGTGACCACTTGTCTGCTGCTCTACCTGCTCTGAG 240
QY 241 TTTTCCCTCTGG 300
DB 241 TTTTCCCTCTGG 300
QY 301 GCTGTGATTTGCTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCGAGCTGGGACG 360
DB 301 GCTGTGATTTGCTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCGAGCTGGGACG 360
QY 361 CCCCTCCGCTCTCAGCCAGGATGCCAAGAGGGCTGGGCCCGGGACCCAGCTGTATCA 420
DB 361 CCCCTCCGCTCTCAGCCAGGATGCCAAGAGGGCTGGGCCCGGGACCCAGCTGTATCA 420
QY 421 CCAAAATCCGAGCACACCCCGCCCAAGCGAGGAGCTGGGCTTCCGAAAGGGAGCTGG 480
DB 421 CCAAAATCCGAGCACACCCCGCCCAAGCGAGGAGCTGGGCTTCCGAAAGGGAGCTGG 480
QY 481 TCACATCTGTGAGGCGCTGTGAGAACTGTGTCACCGCTCAACACACACACAGTG 540
DB 481 TCACATCTGTGAGGCGCTGTGAGAACTGTGTCACCGCTCAACACACACACAGTG 540
QY 541 GACAGAGGGGGCTGTGTGAGCTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 600
DB 541 GACAGAGGGGGCTGTGTGAGCTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 600
QY 601 CCAAGCTCAGCCTCATGCGCTGGTTCACGGGAAAGATCTGGGCGCAGAGAGCTGTCCAG 660
DB 601 CCAAGCTCAGCCTCATGCGCTGGTTCACGGGAAAGATCTGGGCGCAGAGAGCTGTCCAG 660
QY 661 AGCTGACGCTCCCGAGAGATGGGCTGTCTGTGGGGAGTCCGGCGCCACCCGGGG 720
DB 661 AGCTGACGCTCCCGAGAGATGGGCTGTCTGTGGGGAGTCCGGCGCCACCCGGGG 720
QY 721 ACTACGCTCTGTGCTGTGAGCTTTGGCGGAGCTCATCTACCGGCTGTCTACCGGG 780
DB 721 ACTACGCTCTGTGCTGTGAGCTTTGGCGGAGCTCATCTACCGGCTGTCTACCGGG 780
QY 781 AGGGCCACCTCACAATGATGAGCGGTGTCTTTCACACCTCATGAGATGTGTGAGC 840
DB 781 AGGGCCACCTCACAATGATGAGCGGTGTCTTTCACACCTCATGAGATGTGTGAGC 840
QY 841 ATTACAGCAAGAGCAAGGGGCTATCTGTACCAAGCTGTGTGAGACCAAGGAAACAG 900
DB 841 ATTACAGCAAGAGCAAGGGGCTATCTGTACCAAGCTGTGTGAGACCAAGGAAACAG 900
QY 901 GGAACCAAGTGGCGGAGAGAGAGCTGGCCAGGGCGGGCTGGTTACTGAACCTGACATTT 960
DB 901 GGAACCAAGTGGCGGAGAGAGAGCTGGCCAGGGCGGGCTGGTTACTGAACCTGACATTT 960

QY 961 TGACATTGGAGACACAGATCGAGAGGAGAGATTGTGAGCTGTCTGACAGGCTGATACC 1020
DB 961 TGACATTGGAGACACAGATCGAGAGGAGAGATTGTGAGCTGTCTGACAGGCTGATACC 1020
QY 1021 TGGGGCAAAAGTGGCCGTGAAGATATTCAAAGTGTGATGTGACAGCCAGGCTTCTG 1080
DB 1021 TGGGGCAAAAGTGGCCGTGAAGATATTCAAAGTGTGATGTGACAGCCAGGCTTCTG 1080
QY 1081 ACAGAGCGCGCTCATGACCAAGATGACACAGCAAGCTGGTGGCTCTCTGGCGTGA 1140
DB 1081 ACAGAGCGCGCTCATGACCAAGATGACACAGCAAGCTGGTGGCTCTCTGGCGTGA 1140
QY 1141 TCCTGACCAAGGGGCTGTACATTGTGATGAGACAGCTGAGCAAGGGCACTGTGAAT 1200
DB 1141 TCCTGACCAAGGGGCTGTACATTGTGATGAGACAGCTGAGCAAGGGCACTGTGAAT 1200
QY 1201 TTTCTGGGACCGGGGGTGTGAGCCCTGTGAAACCCGCTACGCTGTGCAATTTCTCTGC 1260
DB 1201 TTTCTGGGACCGGGGGTGTGAGCCCTGTGAAACCCGCTACGCTGTGCAATTTCTCTGC 1260
QY 1261 ACGTGGCGGAGGCGATGAGTACCTGTGAGAGCAAGAGCTTGTGCAACCGGACCTGGCG 1320
DB 1261 ACGTGGCGGAGGCGATGAGTACCTGTGAGAGCAAGAGCTTGTGCAACCGGACCTGGCG 1320
QY 1321 CCGGCAACATCTGTGTCTGAGAGAGCTGTGGCCCAAGTCAACGACTTTGGCTGGCCA 1380
DB 1321 CCGGCAACATCTGTGTCTGAGAGAGCTGTGGCCCAAGTCAACGACTTTGGCTGGCCA 1380
QY 1381 AAGCCGAGGGAAGGGGCTGAGCTCAAGACCGGGTGGCCGCTCAAGTGAAGGCGCGGAG 1440
DB 1381 AAGCCGAGGGAAGGGGCTGAGCTCAAGACCGGGTGGCCGCTCAAGTGAAGGCGCGGAG 1440
QY 1441 CTCTCAAAACACGGGAAGTTTACACAGCAAGTGGGATTTTGGGGTCTGCTCT 1500
DB 1441 CTCTCAAAACACGGGAAGTTTACACAGCAAGTGGGATTTTGGGGTCTGCTCT 1500
QY 1501 GGGAGGCTTTCTATATGAGCGGGCTTCGTAACCTTAATATGTACTGAAAGAGGTCCG 1560
DB 1501 GGGAGGCTTTCTATATGAGCGGGCTTCGTAACCTTAATATGTACTGAAAGAGGTCCG 1560
QY 1561 AGGCGGTGAGGAAGGGGTACCGCATGGAACCCCGGAGGGCTTCCAGGGCCCGGTGACG 1620
DB 1561 AGGCGGTGAGGAAGGGGTACCGCATGGAACCCCGGAGGGCTTCCAGGGCCCGGTGACG 1620
QY 1621 TCCTCATGAGCACTGTGTGAGGAGAGCCCGCGGCAACCTTCCGCAAACTGG 1680
DB 1621 TCCTCATGAGCACTGTGTGAGGAGAGCCCGCGGCAACCTTCCGCAAACTGG 1680
QY 1681 CCGAGAAAGTGGCGCCCGGAGCTACGAGTGCAGAGTCCCGACGCTTCCGCAAACTGG 1740
DB 1681 CCGAGAAAGTGGCGCCCGGAGCTACGAGTGCAGAGTCCCGACGCTTCCGCAAACTGG 1740
QY 1741 ACGCCAGAGGCTCACCTCGCCCGAAGCGAGGCTGTGACCCACCGGCTGGGGCCCT 1800
DB 1741 ACGCCAGAGGCTCACCTCGCCCGAAGCGAGGCTGTGACCCACCGGCTGGGGCCCT 1800
QY 1801 TGGCCCCAGAGACCGAGAGAGTGAAGTGGGCTGTGGGGGCACTGTACAGGCCCAAG 1860
DB 1801 TGGCCCCAGAGACCGAGAGAGTGAAGTGGGCTGTGGGGGCACTGTACAGGCCCAAG 1860
QY 1861 AGGGTCCAGGGGGGCAAGTATCTCTGTGTGCGCACAGAGGGGCTGGGCCCAAGTGG 1920
DB 1861 AGGGTCCAGGGGGGCAAGTATCTCTGTGTGCGCACAGAGGGGCTGGGCCCAAGTGG 1920
QY 1921 GGGCTGTGGGGCGGCTGTGACACCCAGACTGTGCAAGATGATCCCGGATTAAGACGG 1980
DB 1921 GGGCTGTGGGGCGGCTGTGACACCCAGACTGTGCAAGATGATCCCGGATTAAGACGG 1980
QY 1981 ATTCTAAGACTCTAAAAA 2000
DB 1981 ATTCTAAGACTCTAAAAA 2000

RESULT 2
LOCUS AR274487 2000 bp DNA Linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6506578.
ACCESSION AR274487
VERSION AR274487.1 GI:29706942
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2000)
AUTHORS Ullrich,A., Gishizky,M. and Sures,I.G.
TITLE Nucleotide encoding megakaryocytic protein tyrosine kinases
JOURNAL Patent: US 6506578-A 1 14-JAN-2003;
FEATURES
source location/Qualifiers
BASE COUNT 375 a 632 c 673 g 320 t
ORIGIN
Query Match 100.0% Score 2000; DB 6; Length 2000;
Best Local Similarity 100.0% Pred. No. 3.2e-289;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGCTCCAGATTGTGACGCGGACCGCTCTGGGGTGTGACGCGGCTCGGAGAGCC 60
DB 1 CTCGCTCCAGATTGTGACGCGGACCGCTCTGGGGTGTGACGCGGCTCGGAGAGCC 60
QY 61 TCCTGG 120
DB 61 TCCTGG 120
QY 121 AGGCTGGGTCCAGTGGGACCGGACCTCCCTGCTGTGGCCAGCGCGGCGGTGGGA 180
DB 121 AGGCTGGGTCCAGTGGGACCGGACCTCCCTGCTGTGGCCAGCGCGGCGGTGGGA 180
QY 181 GGCATTTCCAGCGCTCCCGGACCTGTGACCACTTGTGCTGCTGCTGCTGCTGCTG 240
DB 181 GGCATTTCCAGCGCTCCCGGACCTGTGACCACTTGTGCTGCTGCTGCTGCTGCTG 240
QY 241 TTTCCCTCTGG 300
DB 241 TTTCCCTCTGG 300
QY 301 GCTGTGATTCTGTGAGGAATTTCCCGGGGTGAGCCCGCGTTCCTCCGAGCGTGGCAC 360
DB 301 GCTGTGATTCTGTGAGGAATTTCCCGGGGTGAGCCCGCGTTCCTCCGAGCGTGGCAC 360
QY 361 CCCCTCCGCTGCACGACCGAGATCCACAGAGGCGCTGGGCCCGGCGACCCAGTATCA 420
DB 361 CCCCTCCGCTGCACGACCGAGATCCACAGAGGCGCTGGGCCCGGCGACCCAGTATCA 420
QY 421 CCAATTTGGGACACCGGCGCGCAAGCCAGGAGGAGCTGGCTTCCGCAAGGCGACGTGG 480
DB 421 CCAATTTGGGACACCGGCGCGCAAGCCAGGAGGAGCTGGCTTCCGCAAGGCGACGTGG 480
QY 481 TCACATCTGTGAGAGCTGTGCGAACAAGAGCTGTACCGGCTCAAGCAACCAACAGAG 540
DB 481 TCACATCTGTGAGAGCTGTGCGAACAAGAGCTGTACCGGCTCAAGCAACCAACAGAG 540
QY 541 GACAGAGAGGGGCTGTGAGAGCTGTGCGGCGCTGCGGAGCGGGAGCGCCCTTCGAGACC 600
DB 541 GACAGAGAGGGGCTGTGAGAGCTGTGCGGCGCTGCGGAGCGGGAGCGCCCTTCGAGACC 600
QY 601 CCAAGCTAGCCTCATGCGCGGTGTCCACGGGGAATATCTGGGGCCAGAGAGCTGTCCAGC 660
DB 601 CCAAGCTAGCCTCATGCGCGGTGTCCACGGGGAATATCTGGGGCCAGAGAGCTGTCCAGC 660
QY 661 AGCTCAGACCTCCAGAGAGGGGCTGTCTGTGCGGAGATCCGCGGCGCCAGCCCGGGG 720
DB 661 AGCTCAGACCTCCAGAGAGGGGCTGTCTGTGCGGAGATCCGCGGCGCCAGCCCGGGG 720
QY 721 ACTAGCTCTGTGTGAGCTTTGGCCGCGAGCGTATCCACTACCGCGTGTGACCGCG 780

DB 721 ACTAGCTCTGTGTGAGCTTTGGCCGCGAGCGTATCCACTACCGCGTGTGACCGCG 780
QY 781 ACGCCACCTCAACATTCATGAGCGCGGTCTTCTGCAACCTCATGACATGTGAGAC 840
DB 781 ACGCCACCTCAACATTCATGAGCGCGGTCTTCTGCAACCTCATGACATGTGAGAC 840
QY 841 ATTACGAGAGGACGAAGGGCGTATCTGACCAACCTGTGAGACCAAGGGAACACG 900
DB 841 ATTACGAGAGGACGAAGGGCGTATCTGACCAACCTGTGAGACCAAGGGAACACG 900
QY 901 GGACCAAGTCGCGCCGAGAGAGAGCTGCGCAGGGCGGGTGTGTTACTGAACTGCAGATT 960
DB 901 GGACCAAGTCGCGCCGAGAGAGAGCTGCGCAGGGCGGGTGTGTTACTGAACTGCAGATT 960
QY 961 TGACATTGGGAGACAGATCGGAGAGGAGAGTTTGAAGCTGTCTGCAAGGTTAGTACC 1020
DB 961 TGACATTGGGAGACAGATCGGAGAGGAGAGTTTGAAGCTGTCTGCAAGGTTAGTACC 1020
QY 1021 TGGGCGAAGAGGTGGCCGCGAAGATATCAAGTGTGATGATGACCGCAGCGCTTCG 1080
DB 1021 TGGGCGAAGAGGTGGCCGCGAAGATATCAAGTGTGATGATGACCGCAGCGCTTCG 1080
QY 1081 ACGAGACGCGCTCATGACGAAGATGCAACACGAAACCTGGTGTCTGCTGGGCGTGA 1140
DB 1081 ACGAGACGCGCTCATGACGAAGATGCAACACGAAACCTGGTGTCTGCTGGGCGTGA 1140
QY 1141 TCCTCACCAGGGGCTGTATCATTTGATGAGACACGTAAGCAAGGCAACCTGTGAACT 1200
DB 1141 TCCTCACCAGGGGCTGTATCATTTGATGAGACACGTAAGCAAGGCAACCTGTGAACT 1200
QY 1201 TTTCTGGAGACCGGGGGTGTGAGCGCTGTGAACACCGCTCAGTCTGTGAGTTTCTG 1260
DB 1201 TTTCTGGAGACCGGGGGTGTGAGCGCTGTGAACACCGCTCAGTCTGTGAGTTTCTG 1260
QY 1261 ACGTGGCCGAGGGCATGAGATCTCTGAGAGCAAGAACTGTGACCGCGACCTGGCG 1320
DB 1261 ACGTGGCCGAGGGCATGAGATCTCTGAGAGCAAGAACTGTGACCGCGACCTGGCG 1320
QY 1321 CCCGCAATCTGTGTCACAGAGACCTGTGTGCGCAAGTCAAGCACTTTGGGCTGTGCA 1380
DB 1321 CCCGCAATCTGTGTCACAGAGACCTGTGTGCGCAAGTCAAGCACTTTGGGCTGTGCA 1380
QY 1381 AAGCGGACGGGAGGGGCTAGACTCAACCGGCTGCGTGAAGTGAAGGAGCGCGGAG 1440
DB 1381 AAGCGGACGGGAGGGGCTAGACTCAACCGGCTGCGTGAAGTGAAGGAGCGCGGAG 1440
QY 1441 CTCTCAACACAGGGAATTCAACCAAGTGGATGTGAGTTTGGGGTGTGCTCT 1500
DB 1441 CTCTCAACACAGGGAATTCAACCAAGTGGATGTGAGTTTGGGGTGTGCTCT 1500
QY 1501 GGGAGGTTTCTCATATGAGAGGGGCTCCGTACCCCTAAATGTCACTGAAAGAGGTGTGG 1560
DB 1501 GGGAGGTTTCTCATATGAGAGGGGCTCCGTACCCCTAAATGTCACTGAAAGAGGTGTGG 1560
QY 1561 AGCGCGTGGAGAGGGGATACCGCATGGAACCCCGGAGGGGCTGCAGAGCCCGGTGACG 1620
DB 1561 AGCGCGTGGAGAGGGGATACCGCATGGAACCCCGGAGGGGCTGCAGAGCCCGGTGACG 1620
QY 1621 TCCTCATAGAGCAGTGTGAGAGGACCGCGCGGCGGCAACCTTCCGCAAACTGG 1680
DB 1621 TCCTCATAGAGCAGTGTGAGAGGACCGCGCGGCGGCAACCTTCCGCAAACTGG 1680
QY 1681 CCGAGAACTGTGCGCGGAGAGCTAGCAGATGCAAGTGTGCGGACCTCGCTTCAGGGCAG 1740
DB 1681 CCGAGAACTGTGCGCGGAGAGCTAGCAGATGCAAGTGTGCGGACCTCGCTTCAGGGCAG 1740
QY 1741 ACGCGAGGCTTCCACTTCGCGCCGCAAGCAGAGAGCTTGAACCCAGCCGCTGGGCGCT 1800
DB 1741 ACGCGAGGCTTCCACTTCGCGCCGCAAGCAGAGAGCTTGAACCCAGCCGCTGGGCGCT 1800
QY 1801 TGGCCCAAGAGACCGAGAGAGTGAAGTGTGCGGGGCTGACGACGAGCCCAAG 1860

Db 1801 TGGCCCCAGAGACCGAGAGATGGAGATGGCGGCGTGGGGGCACTGACCAAGGCCCAAG 1860

QY 1861 AGGGTCACAGCGGGCAAGTATCTCTGCTGGCCACAGCAGGAGGCTGGCCACGTAAGG 1920

Db 1861 AGGGTCACAGCGGGCAAGTATCTCTGCTGGCCACAGCAGGAGGCTGGCCACGTAAGG 1920

QY 1921 GGCTCTGGGCGGCGGCTGGACACCCAGACCTGGCAGAGATGATCGCCGATAAAGACG 1980

Db 1921 GGCTCTGGGCGGCGGCTGGACACCCAGACCTGGCAGAGATGATCGCCGATAAAGACG 1980

QY 1981 ATTCTAGAGACTCTAAAAA 2000

Db 1981 ATTCTAGAGACTCTAAAAA 2000

RESULT 3
BC000114
LOCUS
DEFINITION
Homo sapiens, megakaryocyte-associated tyrosine kinase, clone
MGC:1708 IMAGE:3350365, mRNA, complete cds.
BC000114
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2180)
Strausberg, R.
Direct Submission
Submitted (03-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzyzinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saedi, Jacqueline
Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: 1 Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4505108.
Location/Qualifiers
1. 2180
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:4145"
/db_xref="taxon:9606"
/clone="MGC:1708 IMAGE:3350365"
/issue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
400..1923
/codon_start=1
/product="megakaryocyte-associated tyrosine kinase"
/protein_id="AAH00114.1"
/db_xref="GI:12652729"

/translation="MAGRSLSVSWRFHGCDSAEELRVSPRELRAMPBPVSARMT
RRMAPGCTCTKCEHTRPKRGELAFKRGDVITLLECEKNSWYRVNHTSGDGLAA
GARBERALSDADLPKILMPWFHGIISOGEVVOLOPEEDLFLVRSRARGYVLCV
SFGROVYRVLRHGDHITIDEAFCEINLMDEYHNSKRGALCTCTLVPRKHGKRS
ABEELRACHGRLNLQHLITGAOLGEBEFNCAVLOGEYLGOKVANKNKICVYRQAPDE
TAVTAKOHENLRLVGLIHQGLTYMEHVSNGNLVNLRLTRGRALVNTAOLQSL
HVAEGMEYLESKRLVLRDLAARNILVSEDLVAVPSFGLAKARKLIDSLRPVKT
PEALKHGKFTSKSDVMSFGLMEVSEYGRAPYPKMSLEVSAREVGRMEPECCP
GPVHVLMSQWEAEPPARPPFRFLAKRLARELSAGAPASVSQDADGTSRSGP"
BASE COUNT 436 a 689 c 724 g 331 t
ORIGIN
Query Match 98.6%; Score 1971.2; DB 9; Length 2180;
Best Local Similarity 99.7%; Pred. No. 6,3e-285;
Matches 1996; Conservative 0; Mismatches 3; Indels 3; Gaps 2;
QY 1 CTGCTCCAGTGTGTGACGGCGGGACCGCTCGGGGTGTGCACCGCGCTCGCGAGGCC 60
Db 142 CTGCTCCAGTGTGTGACGGCGGGACCGCTCGGGGTGTGCACCGCGCTCGCGAGGCC 201
QY 61 TCCTGGGGGGGGGGGGGGG--GCGGCTCGGGGGGGCGCCCTTGAGAGAAAAAGAGAA 118
Db 202 TCCTGG 261
QY 119 CAGAGCTCGGTCCAGTGGACCCAGCTCCCTACTCTGTGTCAGCGCGGCTGTGG 178
Db 262 CAGAGCTCGGTCCAGTGGACCCAGCTCCCTACTCTGTGTCAGCGCGGCTGTGG 321
QY 179 CAGGCAATCCAGCGTCCCGCGACGTCGACATTCGTCAGTGGCTCGACCTGCTC 238
Db 322 CAGGCAATCCAGCGTCCCGCGACGTCGACATTCGTCAGTGGCTCGACCTGCTC 381
QY 239 AGTTTCCCTTGAGGGGGCGATGGCGGGGGGAGGCTGTGTTTCTGGGGGCAATTTGA 298
Db 382 AGTTTCCCTTGAGGGGGCGATGGCGGGGGGAGGCTGTGTTTCTGGGGGCAATTTGA 440
QY 299 CGGCTGTGATTGTGTGAGGAACTTCCCGGGGTGAGACCCCGCTTCTCGAGCTGCA 358
Db 441 CGGCTGTGATTGTGTGAGGAACTTCCCGGGGTGAGACCCCGCTTCTCGAGCTGCA 500
QY 359 CCGCCCTCCCGTCCAGCGATGGCAAGAGGCGTGGGCGCGGGGACCCAGTGTAT 418
Db 501 CCGCCCTCCCGTCCAGCGATGGCAAGAGGCGTGGGCGCGGGGACCCAGTGTAT 560
QY 419 CACCAATGGAGACACACCCGCCCAAGCGAGGGAGCTGGCTTCGCAAGGCGACGT 478
Db 561 CACCAATGGAGACACACCCGCCCAAGCGAGGGAGCTGGCTTCGCAAGGCGACGT 620
QY 479 GGTACCATCTCTGGAGGCTGCGAGAACAAAGAGCTGTACCGGTCAAGCACACAC 538
Db 621 GGTACCATCTCTGGAGGCTGCGAGAACAAAGAGCTGTACCGGTCAAGCACACAC 680
QY 539 TGGACAGAGAGGGCTGTGTGACGTGGGGGCTGGCGGAGCGGGAGGCCCTCCACAG 598
Db 681 TGGACAGAGAGGGCTGTGTGACGTGGGGGCTGGCGGAGCGGGAGGCCCTCCACAG 740
QY 599 CCCCAGCTCAGCTCATGCCGTGTTTCCACGAGGAAGATTCGCGGCGAGAGGCTGTCA 658
Db 741 CCCCAGCTCAGCTCATGCCGTGTTTCCACGAGGAAGATTCGCGGCGAGAGGCTGTCA 800
QY 659 GCAAGCTCAGCTCCCGAGAGATGGCTGTTCCTGTGGGGAGTCCCGCGCACCCCG 718
Db 801 GCAAGCTCAGCTCCCGAGAGATGGCTGTTCCTGTGGGGAGTCCCGCGCACCCCG 860
QY 719 CGACTACGTCCTGTGGTGTAGCTTGGCGGAGCGATCCACTACCGGCTGTGACCG 778
Db 861 CGACTACGTCCTGTGGTGTAGCTTGGCGGAGCGATCCACTACCGGCTGTGACCG 920
QY 779 CGAGGCGACCTCAGATGAGGCGGTGTCTTCTGCAACCTCATGAGACATGTTGA 838
Db 921 CGAGGCGACCTCAGATGAGGCGGTGTCTTCTGCAACCTCATGAGACATGTTGA 980
QY 839 GCATTACAGCAAGAGCGGCTATCTGTACCAAGCTGTGTGAGCAAAAGCGGAACA 898

Db	981	GCATTTACGACGAGACAGAGGGCGCTATCTGCACCAAGCTGGTAGACCAAGCGGAACA	1040
QY	899	CGGACCAACAAGTCGGCCGAGAGAGGAGACTGGCCAGGGCGGTGTTACTGAACCTGCAGCA	958
Db	1041	CGGAGCAACAAGTCGGCCGAGAGAGGAGCTGGCCAGGGCGGTGTTACTGAACCTGCAGCA	1100
QY	959	TTTGGACATTTGGGAGCACAGATCGGAGAGGGAGATTTGGAGCTCTCTGCAGGGTAGTA	1018
Db	1101	TTTTGACATTTGGGAGCACAGATCGGAGAGGGAGATTTGGAGCTCTCTGCAGGGTAGTA	1160
QY	1019	CCCTGGGCGAAAGGTGGCCGTAAGAGATATCAAGTGTGATGTGCACCCAGGCGTTCTCT	1078
Db	1161	CTTGGGGCGAAAGGTGGCCGTAAGAGATATCAAGTGTGATGTGCACCCAGGCGTTCTCT	1220
QY	1079	GGAGGAGACGGCCGTCATGACGAGAGATGCACAACAGAAACCTGGTGGCTCTCTGGGCGT	1138
Db	1221	GGAGGAGACGGCCGTCATGACGAGAGATGCACAACAGAAACCTGGTGGCTCTCTGGGCGT	1280
QY	1139	GATCTCTGCACAGGGGGCTGTACATTTGTCAATGAGACAGTGAAGCAAGGCGCAACTGTGTAA	1198
Db	1281	GATCTCTGCACAGGGGGCTGTACATTTGTCAATGAGACAGTGAAGCAAGGCGCAACTGTGTAA	1340
QY	1199	CTTTCCTCGAGACCCGGGGGTGAGAGCCCTGTAAGAACCGCTCAGCTCCTGCGAGTTTCTCT	1258
Db	1341	CTTTCCTCGAGACCCGGGGGTGAGAGCCCTGTAAGAACCGCTCAGCTCCTGCGAGTTTCTCT	1400
QY	1359	GCACGTGGCCGAGGGCATGAGTACCTGGAGAGCAAGAAGCTTGTGACACGGCAGCTGGC	1318
Db	1401	GCACGTGGCCGAGGGCATGAGTACCTGGAGAGCAAGAAGCTTGTGACACGGCAGCTGGC	1460
QY	1319	CGCCCGCAACATCTCTGCTCTCAGAGACCTGTGTGGCCAAAGGTGACGCACTTTGGCTTGGC	1378
Db	1461	CGCCCGCAACATCTCTGCTCTCAGAGACCTGTGTGGCCAAAGGTGACGCACTTTGGCTTGGC	1520
QY	1379	CAAGCCGAGCGGAGAGGGGCTAGACTCAAGCCGGCTGGCCCTCAAGTGGAGAGGGCGCCGA	1438
Db	1521	CAAGCCGAGCGGAGAGGGGCTAGACTCAAGCCGGCTGGCCCTCAAGTGGAGAGGGCGCCGA	1580
QY	1439	GGCTCTCAAAACACGGGAAGTTCAACACAGCAAGTCGGATGTCTGGAATTTTGGGGTGTGCT	1498
Db	1581	GGCTCTCAAAACACGGGAAGTTCAACACAGCAAGTCGGATGTCTGGAATTTTGGGGTGTGCT	1640
QY	1499	CTGGGAGGCTCTTCTCATATGAGACGGGCTCCGTACCCATAAATGTCACTGAAGAAGGTGTC	1558
Db	1641	CTGGGAGGCTCTTCTCATATGAGACGGGCTCCGTACCCATAAATGTCACTGAAGAAGGTGTC	1700
QY	1559	GGAGGCGCGTGAGAGAGGGGTACCCCATGGAAGCCGCCGAGGGGCTGTCCAGGCCCCGTGCA	1618
Db	1701	GGAGGCGCGTGAGAGAGGGGTACCCCATGGAAGCCGCCGAGGGGCTGTCCAGGCCCCGTGCA	1760
QY	1619	CGTCCCTCATGAGACCTGTGGGAGGACGAGACCCGCCGCCGCCACCCCTTCCGCACAAT	1678
Db	1761	CGTCCCTCATGAGACCTGTGGGAGGACGAGACCCGCCGCCGCCACCCCTTCCGCACAAT	1820
QY	1679	GGCCGAGAAAGCTGGCCCGGAGAGCTACGACAGTGTGACAGTGTGCCACGCTCTCTCAGGGCA	1738
Db	1821	GGCCGAGAAAGCTGGCCCGGAGAGCTACGACAGTGTGACAGTGTGCCACGCTCTCTCAGGGCA	1880
QY	1739	GGAGCGCAGAGGCTTCACCTCGCCCGGAAGCCAGAGAGCCGTGACCCCAACCCGGTGGGGCC	1798
Db	1881	GGAGCGCAGAGGCTTCACCTCGCCCGGAAGCCAGAGAGCCGTGACCCCAACCCGGTGGGGCC	1940
QY	1799	CTTGGGCCCAAGAGACCGAGAGAGTGTGAGATGTGGGCTGTGGGGCACTAGACAGAGCCCAA	1858
Db	1941	CTTGGGCCCAAGAGACCGAGAGAGTGTGAGAGTGTGGGCTGTGGGGCACTAGACAGAGCCCAA	2000
QY	1859	GGAGGGTCTCAGAGCGGGGCAATCTATCTCTGTGTGCCCAACAGAGGGCTGGCCCACTGTAG	1918
Db	2001	GGAGGGTCTCAGAGCGGGGCAATCTATCTCTGTGTGCCCAACAGAGGGCTGGCCCACTGTAG	2060
QY	1919	GGGGCTCTGGGGCGGGCCGTGGACACCCCAAGACCTGTGGAAGATGATCGCCCGATTAAGAC	1978

DB	2061	GGGGCTCTGGGGGGCCCGTGGACACCCCAAGACCTGCGAAGATGATGCGCCCGATAAAGAC	2120
OY	1979	GGATTCTAAGACCTCTAAAAA	2000
LOCUS	BC003109		
DEFINITION	2121	GGATTCTAAGACCAAAAAA	2142
ACCESSION	BC003109		
KEYWORDS	BC003109.1	GI:13111882	
VERSION	MGC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2180)		
TITLE	Strausberg, R.		
JOURNAL	Direct Submission		
REMARK	Submitted (13-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
COMMENT	NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Ffeli, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywnski, Beta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candace McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeei, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyerdun, Marco Marra.		
FEATURES	Location/Qualifiers		
SOURCE	1..2180		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="LocustID:4145"		
	/db_xref="taxon:9606"		
	/clone="MGC:2101 IMAGE:3357648"		
	/tissue_type="Eye, retinoblastoma"		
	/clone_id="NIH_MGC_16"		
	/lab_host="DH10B-R"		
	/note="Vector: pOTB7"		
	400..1923		
	/codon_start=1		
	/product="megakaryocyte-associated tyrosine kinase"		
	/protein_id="AAH03109.1"		
	/db_xref="GI:13111883"		
	/translation="MAGRSLVSMRAHFHGCDSAEELPRVSPRLRAHPPVSAAMP RRMAGTCTTCCEHTPRKPELFAIRKDDVYITLACNKNMYRKHTSCOEGLLA GALRRREALSDPKLTPMPFHGKISGDAVVOOLQPPEDGILFVRESRHGDDVYLCA SFGRIIVHYRVLNRHGGTLIDAEAFECNLMIMVSHYSSDKDAGICTKLVRPRKRGTKD AAEEELVAGMLNLOHTLGAIGEGEGAVLQGEYLGOKVAARNIKCDVYAOAFLEDE TAVMKMOHEMLVRLGIVTLHQGLYIVNEHSGKLVNLFRTGRGALVNTAOLOFS HVABGEMTEKMLVHRLAARNLTVSDIYAKYSDGLAFAERKGLDSSRLPYKMTA PEALHNGFTSKSDVMSFGVLLMEVFTSGRAPHYKMSLKEYSALVEKGYRMEPEPGCE		

BASE COUNT 436 a 689 c 724 g 331 t
ORIGIN GPHVHMSCHMEAPRRPPRKLAEKLEHLSAGAPASVSGODADGSTSPRSEP*

Query Match 98.6%; Score 1971.2; DB 9; Length 2180;
Best Local Similarity 99.7%; Pred. No. 6.3e-285;
Matches 1996; Conservative 0; Mismatches 3; Indels 3; Gaps 2;

QY 1 CTCGCTCCAAAGTTGTGCAAGCCGGGAGCCGCTGGGGGTGTGAGCCGGCTCGCGAGAGCC 60
Db CTCCCTCCAAAGTTGTGCAAGCCGGGAGCCGCTGGGGGTGTGAGCCGGCTCGCGAGAGCC 201
QY 61 TCCGCGGG 118
Db TCCGCGGG 261
QY 202 TCCGCGGG 178
Db TCCGCGGG 321
QY 119 CCAAGCTCGGTGCAAGTGGCAAGCCAGCTCCCTACCTGTCAGCCAGCCGCTGGCTGTGG 178
Db CCAAGCTCGGTGCAAGTGGCAAGCCAGCTCCCTACCTGTCAGCCAGCCGCTGGCTGTGG 321
QY 179 CAGGCTATCCCAAGCTCCCGAGCTGTGACCACTTGTCTGAGTGTGCTTCCCTGCTGCTC 238
Db CAGGCTATCCCAAGCTCCCGAGCTGTGACCACTTGTCTGAGTGTGCTTCCCTGCTGCTC 381
QY 239 AGTTTCCCTGTGG 298
Db AGTTTCCCTGTGG 440
QY 382 AGTTTCCCTGTGG 440
Db AGTTTCCCTGTGG 500
QY 299 CGGCTGTGATTTGCTGAGGAAGTCCCGGGGTGAGCCCGCTGCTTCCCTGAGAGCTGGA 358
Db CGGCTGTGATTTGCTGAGGAAGTCCCGGGGTGAGCCCGCTGCTTCCCTGAGAGCTGGA 500
QY 441 CGGCTGTGATTTGCTGAGGAAGTCCCGGGGTGAGCCCGCTGCTTCCCTGAGAGCTGGA 500
Db CGGCTGTGATTTGCTGAGGAAGTCCCGGGGTGAGCCCGCTGCTTCCCTGAGAGCTGGA 500
QY 359 CCCCCCTCCGCTCTCAAGCAGAGATGCCAAGAGGCGCTGGGGCCCGGGGCAAGCTGTAT 418
Db CCCCCCTCCGCTCTCAAGCAGAGATGCCAAGAGGCGCTGGGGCCCGGGGCAAGCTGTAT 560
QY 501 CCCCCCTCCGCTCTCAAGCAGAGATGCCAAGAGGCGCTGGGGCCCGGGGCAAGCTGTAT 560
Db CCCCCCTCCGCTCTCAAGCAGAGATGCCAAGAGGCGCTGGGGCCCGGGGCAAGCTGTAT 560
QY 419 CACCAATATGGAGACACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGCGAGCT 478
Db CACCAATATGGAGACACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGCGAGCT 620
QY 561 CACCAATATGGAGACACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGCGAGCT 620
Db CACCAATATGGAGACACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGCGAGCT 680
QY 479 GGTACCATCTGTGAGGCGTCCGAGAACAGAGCTGTGACCGCTCAAGCAGCAGCAGCAG 538
Db GGTACCATCTGTGAGGCGTCCGAGAACAGAGCTGTGACCGCTCAAGCAGCAGCAGCAG 680
QY 621 GGTACCATCTGTGAGGCGTCCGAGAACAGAGCTGTGACCGCTCAAGCAGCAGCAGCAG 680
Db GGTACCATCTGTGAGGCGTCCGAGAACAGAGCTGTGACCGCTCAAGCAGCAGCAGCAG 680
QY 539 TGAACAGAGAGGCGTGTGAGCAGTGGGGCGCTGCGGAGCCGGAGCCCTTCCGAGAG 558
Db TGAACAGAGAGGCGTGTGAGCAGTGGGGCGCTGCGGAGCCGGAGCCCTTCCGAGAG 740
QY 681 TGAACAGAGAGGCGTGTGAGCAGTGGGGCGCTGCGGAGCCGGAGCCCTTCCGAGAG 740
Db TGAACAGAGAGGCGTGTGAGCAGTGGGGCGCTGCGGAGCCGGAGCCCTTCCGAGAG 740
QY 599 CCCCAGAGCTCAGCTCATGCGGTGTTCCACGAGGAAGATCTCGGGCCAGAGAGGCTGTCA 658
Db CCCCAGAGCTCAGCTCATGCGGTGTTCCACGAGGAAGATCTCGGGCCAGAGAGGCTGTCA 800
QY 741 CCCCAGAGCTCAGCTCATGCGGTGTTCCACGAGGAAGATCTCGGGCCAGAGAGGCTGTCA 800
Db CCCCAGAGCTCAGCTCATGCGGTGTTCCACGAGGAAGATCTCGGGCCAGAGAGGCTGTCA 800
QY 659 GCAAGCTCAGAGCTCCGAGAGAGGCGCTGTCTGTGTGCGGAGTCCGCGCCAGCCCGG 718
Db GCAAGCTCAGAGCTCCGAGAGAGGCGCTGTCTGTGTGCGGAGTCCGCGCCAGCCCGG 860
QY 801 GCAAGCTCAGAGCTCCGAGAGAGGCGCTGTCTGTGTGCGGAGTCCGCGCCAGCCCGG 860
Db GCAAGCTCAGAGCTCCGAGAGAGGCGCTGTCTGTGTGCGGAGTCCGCGCCAGCCCGG 860
QY 719 CGAGCTAGTCTGTGAGTGTGAGCTTTGGCCCGAGCTGATCCACTACCGGCTGTGACCG 778
Db CGAGCTAGTCTGTGAGTGTGAGCTTTGGCCCGAGCTGATCCACTACCGGCTGTGACCG 920
QY 861 CGAGCTAGTCTGTGAGTGTGAGCTTTGGCCCGAGCTGATCCACTACCGGCTGTGACCG 920
Db CGAGCTAGTCTGTGAGTGTGAGCTTTGGCCCGAGCTGATCCACTACCGGCTGTGACCG 980
QY 779 CGAGCGGCACTCACAATGATGAGGCGGTCTTCTTGTGCAACCTCATGAGATGAGTGA 838
Db CGAGCGGCACTCACAATGATGAGGCGGTCTTCTTGTGCAACCTCATGAGATGAGTGA 980
QY 921 CGAGCGGCACTCACAATGATGAGGCGGTCTTCTTGTGCAACCTCATGAGATGAGTGA 980
Db CGAGCGGCACTCACAATGATGAGGCGGTCTTCTTGTGCAACCTCATGAGATGAGTGA 1040
QY 839 GCAATTCAGCAAGAGAGAGGCGCTATCTGCAACCAAGCTGTGAGACCAAGGGAACA 898
Db GCAATTCAGCAAGAGAGAGGCGCTATCTGCAACCAAGCTGTGAGACCAAGGGAACA 1040
QY 981 GCAATTCAGCAAGAGAGAGGCGCTATCTGCAACCAAGCTGTGAGACCAAGGGAACA 1040
Db GCAATTCAGCAAGAGAGAGGCGCTATCTGCAACCAAGCTGTGAGACCAAGGGAACA 1040
QY 899 CGGAGCAATGCGCGGAGAGAGAGCTGCGCAGAGGGGGGTGTACTGAGACTGAGCA 958
Db CGGAGCAATGCGCGGAGAGAGAGCTGCGCAGAGGGGGGTGTACTGAGACTGAGCA 1100
QY 1041 CGGAGCAATGCGCGGAGAGAGAGCTGCGCAGAGGGGGGTGTACTGAGACTGAGCA 1100
Db CGGAGCAATGCGCGGAGAGAGAGCTGCGCAGAGGGGGGTGTACTGAGACTGAGCA 1100
QY 959 TTTGACATTTGGAGACAGATCGAGAGAGGAGATTGGAGCTGTCTCGAGGGTGA 1018
Db TTTGACATTTGGAGACAGATCGAGAGAGGAGATTGGAGCTGTCTCGAGGGTGA 1018

Db |||||
QY 1101 TTTGACATTTGGAGACAGATCGAGAGAGGAGATTGGAGCTGTCTCGAGGGTGA 1160
QY 1019 CCTGGGCAAAAGGTGGCCGCTGAAGATATCAAGTGTGATGACAGCCAGCCCTTCT 1078
Db CCTGGGCAAAAGGTGGCCGCTGAAGATATCAAGTGTGATGACAGCCAGCCCTTCT 1220
QY 1161 CCTGGGCAAAAGGTGGCCGCTGAAGATATCAAGTGTGATGACAGCCAGCCCTTCT 1220
Db CCTGGGCAAAAGGTGGCCGCTGAAGATATCAAGTGTGATGACAGCCAGCCCTTCT 1220
QY 1079 GGAGAGAGGCGCTCATGAGAGATGCAACACAGAGAACCTGTGCTGTGGGCGT 1138
Db GGAGAGAGGCGCTCATGAGAGATGCAACACAGAGAACCTGTGCTGTGGGCGT 1280
QY 1221 GGAGAGAGGCGCTCATGAGAGATGCAACACAGAGAACCTGTGCTGTGGGCGT 1280
Db GGAGAGAGGCGCTCATGAGAGATGCAACACAGAGAACCTGTGCTGTGGGCGT 1340
QY 1139 GATCCTGACAGAGGCGCTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
Db GATCCTGACAGAGGCGCTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1340
QY 1199 CTTTCTGCGAGCCGGGGGTGAGCCCTGCGAAGACCGCTGAGCTGCTGAGTTTCTCT 1258
Db CTTTCTGCGAGCCGGGGGTGAGCCCTGCGAAGACCGCTGAGCTGCTGAGTTTCTCT 1400
QY 1341 CTTTCTGCGAGCCGGGGGTGAGCCCTGCGAAGACCGCTGAGCTGCTGAGTTTCTCT 1400
Db CTTTCTGCGAGCCGGGGGTGAGCCCTGCGAAGACCGCTGAGCTGCTGAGTTTCTCT 1400
QY 1259 GCAGGTGGCGAGGCGATGAGATGATGAGAGAGCAGAGAACCTGTGACCGGCACTGGC 1318
Db GCAGGTGGCGAGGCGATGAGATGAGAGAGCAGAGAACCTGTGACCGGCACTGGC 1460
QY 1401 GCAGGTGGCGAGGCGATGAGATGAGAGAGCAGAGAACCTGTGACCGGCACTGGC 1460
Db GCAGGTGGCGAGGCGATGAGATGAGAGAGCAGAGAACCTGTGACCGGCACTGGC 1520
QY 1319 CGCCCGCAACATCTGTGATGAGAGACCTGTGTGAGTGAAGTGAAGTGAAGTGAAGTGA 1378
Db CGCCCGCAACATCTGTGATGAGAGACCTGTGTGAGTGAAGTGAAGTGAAGTGAAGTGA 1520
QY 1461 CGCCCGCAACATCTGTGATGAGAGACCTGTGTGAGTGAAGTGAAGTGAAGTGAAGTGA 1520
Db CGCCCGCAACATCTGTGATGAGAGACCTGTGTGAGTGAAGTGAAGTGAAGTGAAGTGA 1520
QY 1379 CAAAGCCGAGAGAGGAGGCTAGACTCAACCCGCTCCGCTCAAGTGAAGGAGGAGGAGG 1438
Db CAAAGCCGAGAGAGGAGGCTAGACTCAACCCGCTCCGCTCAAGTGAAGGAGGAGGAGG 1580
QY 1521 CAAAGCCGAGAGAGGAGGCTAGACTCAACCCGCTCCGCTCAAGTGAAGGAGGAGGAGG 1580
Db CAAAGCCGAGAGAGGAGGCTAGACTCAACCCGCTCCGCTCAAGTGAAGGAGGAGGAGG 1640
QY 1439 GCGCTCTCAAAACAGGAGATTCACACAGAGTGGATGTGTGAGATTTGGGTGTGCT 1498
Db GCGCTCTCAAAACAGGAGATTCACACAGAGTGGATGTGTGAGATTTGGGTGTGCT 1640
QY 1581 GCGCTCTCAAAACAGGAGATTCACACAGAGTGGATGTGTGAGATTTGGGTGTGCT 1640
Db GCGCTCTCAAAACAGGAGATTCACACAGAGTGGATGTGTGAGATTTGGGTGTGCT 1640
QY 1499 CTGGGAGAGCTTCTCATATGAGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1558
Db CTGGGAGAGCTTCTCATATGAGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
QY 1641 CTGGGAGAGCTTCTCATATGAGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
Db CTGGGAGAGCTTCTCATATGAGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
QY 1559 GGAAGGCGGTGAGAGAGGAGTACCGCATGAGAACCCCGGAGAGGCTGTCCAGAGCCCGTGA 1618
Db GGAAGGCGGTGAGAGAGGAGTACCGCATGAGAACCCCGGAGAGGCTGTCCAGAGCCCGTGA 1760
QY 1701 GGAAGGCGGTGAGAGAGGAGTACCGCATGAGAACCCCGGAGAGGCTGTCCAGAGCCCGTGA 1760
Db GGAAGGCGGTGAGAGAGGAGTACCGCATGAGAACCCCGGAGAGGCTGTCCAGAGCCCGTGA 1760
QY 1619 GGTCTCTATGAGAGCTGTGAGAGGAGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 1678
Db GGTCTCTATGAGAGCTGTGAGAGGAGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 1820
QY 1761 GGTCTCTATGAGAGCTGTGAGAGGAGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 1820
Db GGTCTCTATGAGAGCTGTGAGAGGAGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 1820
QY 1679 GGCAGAGAGCTGTGAGAGGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1738
Db GGCAGAGAGCTGTGAGAGGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1880
QY 1821 GGCAGAGAGCTGTGAGAGGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1880
Db GGCAGAGAGCTGTGAGAGGAGTACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1880
QY 1739 GGAAGGCGAGGAGCTCAGCTGCGCCGAGAACCCAGAGAGCCCTGACCCAGCCGCTGGGGCC 1798
Db GGAAGGCGAGGAGCTCAGCTGCGCCGAGAACCCAGAGAGCCCTGACCCAGCCGCTGGGGCC 1940
QY 1881 GGAAGGCGAGGAGCTCAGCTGCGCCGAGAACCCAGAGAGCCCTGACCCAGCCGCTGGGGCC 1940
Db GGAAGGCGAGGAGCTCAGCTGCGCCGAGAACCCAGAGAGCCCTGACCCAGCCGCTGGGGCC 1940
QY 1799 CTTGGCCCGCAGAGAGCAGAGAGTGGAGAGTGGGCGCTGGGGGCACTGACAGGCCAA 1858
Db CTTGGCCCGCAGAGAGCAGAGAGTGGAGAGTGGGCGCTGGGGGCACTGACAGGCCAA 2000
QY 1941 CTTGGCCCGCAGAGAGCAGAGAGTGGAGAGTGGGCGCTGGGGGCACTGACAGGCCAA 2000
Db CTTGGCCCGCAGAGAGCAGAGAGTGGAGAGTGGGCGCTGGGGGCACTGACAGGCCAA 2000
QY 1859 GGAAGGCTCAGAGGAGGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1918
Db GGAAGGCTCAGAGGAGGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2060
QY 2001 GGAAGGCTCAGAGGAGGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2060
Db GGAAGGCTCAGAGGAGGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2060
QY 1919 GGGGCTGTGGGGCGCCGTGAGACACCCAGAGCTGCGAAGAGATGATGCGCCGTAAGAGAC 1978
Db GGGGCTGTGGGGCGCCGTGAGACACCCAGAGCTGCGAAGAGATGATGCGCCGTAAGAGAC 2120
QY 2061 GGGGCTGTGGGGCGCCGTGAGACACCCAGAGCTGCGAAGAGATGATGCGCCGTAAGAGAC 2120
Db GGGGCTGTGGGGCGCCGTGAGACACCCAGAGCTGCGAAGAGATGATGCGCCGTAAGAGAC 2120
QY 1979 GATTTCTAAGAGATCTTAAAAA 2000
Db GATTTCTAAGAGATCTTAAAAA 2142

RESULT 5

LOCUS	BD176710	1997 bp	DNA	linear	PAT 18-MAR-2003
DEFINITION	Method of examining allergic disease.				
ACCESSION	BD176710				
VERSION	BD176710.1	GI:291222420			
KEYWORDS	MO 02075304 -A/11.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Sugita Y., Heishi M., Kagaya S., Gunji S. and Tsujimoto G.				
TITLE	Method of examining allergic disease				
JOURNAL	Patent: WO 02075304-A 11 26-SEP-2002; GENEX RESEARCH INC., JAPAN AS REPRESENTED BY GENERAL DIRECTOR OF NATIONAL CHILDREN'S HOSPITAL, HIDEOTO YONEKURA, YASUHIKO YAMAOKA, SHIGERU SAKURAI, TAKUO MATANABE YUJI SUGITA, MASAYUKI HEISHI, SHINJI KAGAYA, SHIGEMICHI GUNJI, GOZO TSUJIMOTO				
COMMENT	OS Homo sapiens (human) PN MO 02075304-A/11 PD 26-SEP-2002 PE 01-MAR-2002 WO 2002JP001916 PR 21-MAR-2001 JP 01P 081028 PI YUJI SUGITA, MASAYUKI HEISHI, SHINJI KAGAYA, SHIGEMICHI GUNJI, PI GOZO TSUJIMOTO PC G01N33/15, G01N33/50, C12N15/09, C12Q1/02, C12Q1/68, A61K45/00, PC A61K31/11, PC A61K39/395, A61P37/08, A61K48/00, A01K67/027 CC Method of examining allergic disease FH Key Location/Qualifiers FT CDS (266). (1789).				
FEATURES	Source 1. 1997 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"				
BASE COUNT	372 a 631 c 676 g 318 t				
ORIGIN					
Query Match	98.2%; Score 1963.4; DB 6; Length 1997;				
Best Local Similarity	99.8%; Pred. No. 9.4e-284;				
Matches 1987; Conservative	0; Mismatches 1; Indels 3; Gaps 2;				
QY	1 CTCGGTCCAAGTTGTGCACGCCGGACGCCCTCGGGGTGTGCAGCCGGCTCGCGAGGCC 60				
DB	8 CTCGGTCCAAGTTGTGCACGCCGGACGCCCTCGGGGTGTGCAGCCGGCTCGCGAGGCC 67				
QY	61 TCCGAGGGGGGGGGGGGG--GGGGCTCGGGGGGGGGGGGGGAGCAGAAAACGAGAA 118				
DB	68 TCCGAGGG 127				
QY	119 CCAGGCTCGGTACAGTGTGCACCCAGACGCTCCCTAACCTCGTGTCCAGAGCCCTGTGG 178				
DB	128 CCAGGCTCGGTACAGTGTGCACCCAGACGCTCCCTAACCTCGTGTCCAGAGCCCTGTGG 187				
QY	179 CAGGCGATTCCAGCGCTCCCGGACGTGTGACCACTTGTCTCACTGTGCTCTCACTGTCTC 238				
DB	188 CAGGCGATTCCAGCGCTCCCGGACGTGTGACCACTTGTCTCACTGTGCTCTCACTGTCTC 247				
QY	239 AGTTTCCCTCTGGGGGGGGGATGGGGGGGGGAGGCTCTGTGGTTTCTTGGGGGATTTCA 298				
DB	248 AGTTTCCCTCTGGGGGGGGGATGGGGGGGGGAGGCTCTGTGGTTTCTTGGGGGATTTCA 306				
QY	299 CGGCTGTATTCTCGTGTAGAACTTCCCGGGGTGAGCCCGGCTTCTTCGAGCCTGTGCA 358				
DB	307 CGGCTGTATTCTCGTGTAGAACTTCCCGGGGTGAGCCCGGCTTCTTCGAGCCTGTGCA 366				
QY	359 CCCCCCTCCGCTCAGCCAGAGTCCAAAGAGGCGTGGGGCCCGGGGACCCAGCTGTAT 418				
DB	367 CCCCCCTCCGCTCAGCCAGAGTCCAAAGAGGCGTGGGGCCCGGGGACCCAGCTGTAT 426				
QY	419 CACCAAAATGCAGACACCCGCCCAAGCCAGGGAGTGGCTTCCGCAAGGGGAGGT 478				

D	b	427	CACCAAAATGCGAGCACACCCGCCCAAGCCAGGGGAGCTGGCTTTCGCGAAGGGCGAGCT	486
Q	y	479	GGATACCAATCCTGGAGGCGCTCGAGAAACAAGAGCTGTACCGCGTCAAGCACACACCAAG	538
D	b	487	GGTACCAATCCTGGAGGCGCTCGAGAAACAAGAGCTGTACCGCGTCAAGCACACACCAAG	546
Q	y	539	TGGACAGAGAGGGGCTGCTGGCAGCTGGGGCGCTCGGAGAGGGAGAGGCCCTCTCCGAGA	598
D	b	547	TGGACAGAGAGGGGCTGCTGGCAGCTGGGGCGCTCGGAGAGGGAGAGGCCCTCTCCGAGA	606
Q	y	599	CCCCAAGCTCAGCCCTCATGCCGTGTTCCAGAGGAGATCTCGGGCCAGAGAGCGCTGCCA	658
D	b	607	CCCCAAGCTCAGCCCTCATGCCGTGTTCCAGAGGAGATCTCGGGCCAGAGAGCGCTGCCA	666
Q	y	659	GCAGCTGCAGCCTCCCGAGATGGGCTGTTCCTGTGGGGAGTCCCGCGCCACCCCGG	718
D	b	667	GCAGCTGCAGCCTCCCGAGATGGGCTGTTCCTGTGGGGAGTCCCGCGCCACCCCGG	726
Q	y	719	CGACTACGCTCGTGGCGTGAAGCTTTGGCGCGAGCTCATCCACTACCGCGTGCACACG	778
D	b	727	CGACTACGCTCGTGGCGTGAAGCTTTGGCGCGAGCTCATCCACTACCGCGTGCACACG	786
Q	y	779	CGACGGCCACCTCAATCGATGAGAGCCGTTTCTTGCACAACTCATATGACATGTGTGA	838
D	b	787	CGACGGCCACCTCAATCGATGAGAGCCGTTTCTTGCACAACTCATATGACATGTGTGA	846
Q	y	839	GCATTTCAGCAGAGCAAGAGGGCGCTATCTGCACCAAGCTGTGTGAGACCAAGCGGAACA	898
D	b	847	GCATTTCAGCAGAGCAAGAGGGCGCTATCTGCACCAAGCTGTGTGAGACCAAGCGGAACA	906
Q	y	899	CGGAGCAAGTCCGCGCGAGAGAGAGCTGCCAGGGCGGGCTGTACTGAACCTGACACA	958
D	b	907	CGGAGCAAGTCCGCGCGAGAGAGAGCTGCCAGGGCGGGCTGTACTGAACCTGACACA	966
Q	y	959	TTTTCGACTTGGGAGACAGATCGGAGAGGAGAGTTTGACTCTCTCCGACGGTGAAGTA	1018
D	b	967	TTTTCGACTTGGGAGACAGATCGGAGAGGAGAGTTTGACTCTCTCCGACGGTGAAGTA	1026
Q	y	1019	CCTGGGGCAAAAAGTGGCGCGTGAAGAAATATCAAGTGTGATGTGACAGCCAGGCGTTTCT	1078
D	b	1027	CCTGGGGCAAAAAGTGGCGCGTGAAGAAATATCAAGTGTGATGTGACAGCCAGGCGTTTCT	1086
Q	y	1079	GGAGCAGACGGCGCTCATGACGAGAGATCAACACAGAGAACTGGTGGCTCTTCCTGGGCGT	1138
D	b	1087	GGAGCAGACGGCGCTCATGACGAGAGATCAACACAGAGAACTGGTGGCTCTTCCTGGGCGT	1146
Q	y	1139	GATCCCTGACAGAGGGGCTGTACATTGTATGAGACAGCTGAGACAAAGGCAACCTGTGTAA	1198
D	b	1147	GATCCCTGACAGAGGGGCTGTACATTGTATGAGACAGCTGAGACAAAGGCAACCTGTGTAA	1206
Q	y	1199	CTTTCCTGCGGAGCCCGGGGCTCGAGCCCTGTGAACACCGCTCAGCTCTCGAGTTTCTCT	1258
D	b	1207	CTTTCCTGCGGAGCCCGGGGCTCGAGCCCTGTGAACACCGCTCAGCTCTCGAGTTTCTCT	1266
Q	y	1259	GCAGCTGGCCCGAGGCGATGAGATCTTGGAGAGCAAGAGCTTGTGCACCGCGACCTGCG	1318
D	b	1267	GCAGCTGGCCCGAGGCGCATGAGATCTTGGAGAGCAAGAGCTTGTGCACCGCGACCTGCG	1326
Q	y	1319	CGCCCGGCAAAATCCTGTGCTCAGAGAGACCTGGTGGCCAAAGGCAAGCGACTTTGGGCTGGC	1378
D	b	1327	CGCCCGGCAAAATCCTGTGCTCAGAGAGACCTGGTGGCCAAAGGCAAGCGACTTTGGGCTGGC	1386
Q	y	1379	CAAAAGCCAGAGGGAAGGGGCTAGACTCAAGCGCGGTGCCGCTCAAGTGGAGCGCGCCCGA	1438
D	b	1387	CAAAAGCCAGAGGGAAGGGGCTAGACTCAAGCGCGGTGCCGCTCAAGTGGAGCGCGCCCGA	1446
Q	y	1439	GGCTCTCAACACAGGGAAGTTCCACAGCAAGTCGGATGCTTGGAGTTTGGGGGTCTCT	1498
D	b	1447	GGCTCTCAACACAGGGAAGTTCCACAGCAAGTCGGATGCTTGGAGTTTGGGGGTCTCT	1506
Q	y	1499	CTGGGAGGCTTCTCATATGTGACGGGGCTCGTACCCCTAAATGTGCATCGAAGAGGTTGC	1558
D	b	1507	CTGGGAGGCTTCTCATATGTGACGGGGCTCGTACCCCTAAATGTGCATCGAAGAGGTTGC	1566

QY 1559 GGAAGCCGTGAGAAAGGGGATCCGCATGGAACCCCGAGGGCTGTCCAGGCCCCCTGCA 1618
| | | | |
Db 1567 GGAAGCCGTGAGAAAGGGGATCCGCATGGAACCCCGAGGGCTGTCCAGGCCCCCTGCA 1626
| | | | |
QY 1619 CGTCTCATGAGACGCTGCTGGGAGGAGACCCCGCCCGCCGACACCTTCCGCAACT 1678
| | | | |
Db 1627 CGTCTCATGAGACGCTGCTGGGAGGAGACCCCGCCCGCCGACACCTTCCGCAACT 1686
| | | | |
QY 1679 GCGCAGAGAGCTGGCCCGGAGAGTACGACAGTGCAGTGTCCACCTTCCGCTGAGGCA 1728
| | | | |
Db 1687 GCGCAGAGAGCTGGCCCGGAGAGTACGACAGTGTCCACCTTCCGCTGAGGCA 1746
| | | | |
QY 1739 GGAAGCCGAGCGCTTCACCTTGCCCGGAGAGCCCTGACCCCAACCCGCTGGGACC 1798
| | | | |
Db 1747 GGAAGCCGAGCGCTTCACCTTGCCCGGAGAGCCCTGACCCCAACCCGCTGGGACC 1806
| | | | |
QY 1799 CTTGGCCCGAGAGAGAGAGATGAGAGTGGGGGTGGGGGCTGACATGACAGGCCCA 1858
| | | | |
Db 1807 CTTGGCCCGAGAGAGAGAGATGAGAGTGGGGGTGGGGGCTGACATGACAGGCCCA 1866
| | | | |
QY 1859 GGAAGGTTCAGAGCGGCGCAATCATCTCTGTGTGCCACAGACAGAGGGCTGGCCACGTAG 1918
| | | | |
Db 1867 GGAAGGTTCAGAGCGGCGCAATCATCTCTGTGTGCCACAGACAGAGGGCTGGCCACGTAG 1926
| | | | |
QY 1919 GGGGCTCTGGGCGCCGCTGGACACCCAGACTGCGAAGAGATGCGCCGATTAAGAC 1978
| | | | |
Db 1927 GGGGCTCTGGGCGCCGCTGGACACCCAGACTGCGAAGAGATGCGCCGATTAAGAC 1986
| | | | |
QY 1979 GGATTTCTAAG 1989
| | | | |
Db 1987 GGATTTCTAAG 1997
| | | | |

RESULT 6
AR084686 1987 bp DNA linear PAT 01-SEP-2000
LOCUS Sequence 1 from patent US 5981201.
DEFINITION AR084686
ACCESSION AR084686
VERSION AR084686.1 GI:10011456
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1987)
AUTHORS Avraham, H. and Groopman, J. E.
TITLE Methods of detection and treatment of breast cancer
JOURNAL Patent: US 5981201-A 1 09-NOV-1999;
FEATURES location/Qualifiers
source 1. 1987
BASE COUNT 369 a 628 c 672 g 318 t
ORIGIN

Query Match 95.5%; Score 1909.8; DB 6; Length 1987;
Best Local Similarity 99.2%; Pred. No. 9.5e-276;
Matches 1973; Conservative 0; Mismatches 7; Indels 9; Gaps 5;

QY 241 TTTCCTCTGGGGGCGATGCGGGGAGGCTCTGTGTTTCTGCGGGCATTTTCACG 300
| | | | |
Db 248 TTTCCTCTGGGGGCGATGCGGGGAGGCTCTGTGTTTCTGCGGGCATTTTCACG 305
| | | | |
QY 301 GCTGTGATTTCTGTGAGAACCTTCCCGGTGAGCCCCCGCTTCTCCGACCTGGAC 360
| | | | |
Db 306 GCTGTGATTTCTGTGAGAACCTTCCCGGTGAGCCCCCGCTTCTCCGACCTGGAC 365
| | | | |
QY 361 CCCCTCCGCTCTGAGAGATGCAACAGAGGCGCTGGGCGCCGGGAGCCAGCTATCA 420
| | | | |
Db 366 CCCCTCCGCTCTGAGAGATGCAACAGAGGCGCTGGGCGCCGGGAGCCAGCTATCA 425
| | | | |
QY 421 CCAATGCGAGCACACCCGCCCAAGCAGGGAGTGTGGCTTCCGAAAGGCGACGTGG 480
| | | | |
Db 426 CCAATGCGAGCACACCCGCCCAAGCAGGGAGTGTGGCTTCCGAAAGGCGACGTGG 485
| | | | |
QY 481 TCACCATCTGTGAGGCTTGGAGAACAGAGCTGTACCGGCTCAGCACACACAGTG 540
| | | | |
Db 486 TCACCATCTGTGAGGCTTGGAGAACAGAGCTGTACCGGCTCAGCACACACAGTG 545
| | | | |
QY 541 GACAGAGGGGCTGTGGAGCTGGGGGCGTGGCGGAGCGGGAGGCGCTTCCGACAC 600
| | | | |
Db 546 GACAGAGGGGCTGTGGAGCTGGGGGCGTGGCGGAGCGGGAGGCGCTTCCGACAC 605
| | | | |
QY 601 CCAAGCTCAGCTTCATGCGCTGTGTCCAGGGGAAGATCTCGGGCAGAGAGGCTGTCCAG 660
| | | | |
Db 606 CCAAGCTCAGCTTCATGCGCTGTGTCCAGGGGAAGATCTCGGGCAGAGAGGCTGTCCAG 665
| | | | |
QY 661 AGCTGCAGCTCCCGAGAGATGGGCTGTCTGTGGGAGTCCGCGCCACCCGGCG 720
| | | | |
Db 666 AGCTGCAGCTCCCGAGAGATGGGCTGTCTGTGGGAGTCCGCGCCACCCGGCG 725
| | | | |
QY 721 ACTACGCTGTGGGTGAGCTTTGGCGGAGCGATCATCACTACCGGCTGTGACCGCG 780
| | | | |
Db 726 ACTACGCTGTGGGTGAGCTTTGGCGGAGCGATCATCACTACCGGCTGTGACCGCG 785
| | | | |
QY 781 ACGGCCACTTCACATGATGAGGCGCTGTCTTCTGCAACCTCATGAGCATGTGAGG 840
| | | | |
Db 786 ACGGCCACTTCACATGATGAGGCGCTGTCTTCTGCAACCTCATGAGCATGTGAGG 845
| | | | |
QY 841 ATTACAGCAAGAGCAAGGGGCTATCTGCAACCAAGCTGTGTGAGACCAACGGAACAG 900
| | | | |
Db 846 ATTACAGCAAGAGCAAGGGGCTATCTGCAACCAAGCTGTGTGAGACCAACGGAACAG 905
| | | | |
QY 901 GGAACAAGTGGCGGAGAGAGAGCTGGCCAGGGCGGCTGTTACTGAACCTCAGCAT 960
| | | | |
Db 906 GGAACAAGTGGCGGAGAGAGAGCTGGCCAGGGCGGCTGTTACTGAACCTCAGCAT 965
| | | | |
QY 961 TGACATTTGGAGCAGAGATGGAAGGAGATTTGGAGCTGTCTGCAAGGTGATACC 1020
| | | | |
Db 966 TGACATTTGGAGCAGAGATGGAAGGAGATTTGGAGCTGTCTGCAAGGTGATACC 1025
| | | | |
QY 1021 TGGGGCAAAAGTGGCGGCTGTAAGATATCAAGTGTATGTGACAGCCGAGGCTTCTG 1080
| | | | |
Db 1026 TGGGGCAAAAGTGGCGGCTGTAAGATATCAAGTGTATGTGACAGCCGAGGCTTCTG 1085
| | | | |
QY 1081 ACGAGAGCGCGCTCATGAGCAAGATGCAACACAGAACTGGTGGCTCTCTGGGCTGA 1140
| | | | |
Db 1086 ACGAGAGCGCGCTCATGAGCAAGATGCAACACAGAACTGGTGGCTCTCTGGGCTGA 1145
| | | | |
QY 1141 TCCTGCACAGAGGGCTGTATATTGTATGAGAGCAGTGAAGAGGCAACCTGGTA 1200
| | | | |
Db 1146 TCCTGCACAGAGGGCTGTATATTGTATGAGAGCAGTGAAGAGGCAACCTGGTA 1205
| | | | |
QY 1201 TTCTGGGAGCCCGGGGCTCAAGCCCTGTGAACACCCCTCAGCTCTGAGTTTCTG 1260
| | | | |
Db 1206 TTCTGGGAGCCCGGGGCTCAAGCCCTGTGAACACCCCTCAGCTCTGAGTTTCTG 1265
| | | | |
QY 1261 ACGTGGCGAGGCGATGAGATACCTGAGAGCAAGAGCTGTGACACCGGACCTGGCG 1320
| | | | |
Db 1266 ACGTGGCGAGGCGATGAGATACCTGAGAGCAAGAGCTGTGACACCGGACCTGGCG 1325
| | | | |

QY	1321	CCGGCAACATCTTGCTCACAAGAACCTGGTGGCCAAAGTACAGGACTTTGGCTGGCGCA	1380
QY	1326	CCCCGAACATCTTGCTCACAAGAACCTGGTGGCCAAAGTACAGGACTTTGGCTGGCGCA	1385
QY	1381	AAGCGAGACCGGAAGGGCTGTACATCAAGCCCGCTGCCGTCAAGTGAAGGACGGCCCGAAG	1440
QY	1386	AAGCGAGACCGGAAGGGCTGTACATCAAGCCCGCTGCCGTCAAGTGAAGGACGGCCCGAAG	1445
QY	1441	CTCTCAAAACACGGGAAGTTCCACAGCAAGTGGATGTCTGGAGTTTGGGCTGCTCTCT	1500
QY	1446	CTCTCAAAACACGGG---GTTCAACACAGCAAGTGGATGTCTGGAGTTTGGGCTGCTCTCT	1502
QY	1501	GGGAGGCTTCTCATATGAGACGGGCTCCGTACCTTAAATGTCACTGAAGAAGGTGTCCG	1560
QY	1503	GGGAGGCTTCTCATATGAGACGGGCTCCGTACCTTAAATGTCACTGAAGAAGGTGTCCG	1562
QY	1561	AGGCGGTGGAAGGGGTACCGCAATGGAAACCCCGAGGGGTGTCCAGAGCCCGCTGCACG	1620
QY	1563	AGGCGGTGGAAGGGGTACCGCAATGGAAACCCCGAGGGGTGTCCAGAGCCCGCTGCACG	1622
QY	1621	TCTGTATGACAGCTGTCTGGGAGGACAGAGCCCGCGCCGACACCTTCGCAAACTGG	1680
QY	1623	TCTGTATGACAGCTGTCTGGGAGGACAGAGCCCGCGCCGACACCTTCGCAAACTGG	1681
QY	1681	CCGAGAAAGCTGGCCCGGGAGCTACGACAGTACGAGGTGCCCAAGCTCCGTCCAGGGCAGG	1740
QY	1682	CCGAGAAAGCTGGCCCGGGAGCTACGACAGTACGAGGTGCCCAAGCTCCGTCCAGGGCAGG	1741
QY	1741	ACGGCGAGCGGTCCACCTGGCCCGGAAAGCCAGAGACCTTACCCCACTCCGGTGGGGCCT	1800
QY	1742	ACGGCGAGCGG-TCCACCTGGCCCGGAAAGCCAGAGACCTTACCCCACTCCGGTGGGGCCT	1798
QY	1801	TGGCCCGCAGAGACCGAGAGAGTGGAGAGTCCGCGTGGGGGACACTGCACAGGCCCAAGG	1860
QY	1799	TGGCCCGCAGAGACCGAGAGAGTGGAGAGTCCGCGTGGGGGACACTGCACAGGCCCAAGG	1858
QY	1861	AGGGTCCAGCGCGGCAAGTCACTCTCTGTGTCGCACAGCAGGGGGTGGGCCACGTAGGG	1920
QY	1859	AGGGTCCAGCGCGGCAAGTCACTCTCTGTGTCGCACAGCAGGGGGTGGGCCACGTAGGG	1918
QY	1921	GGGCTTGGGGGGCCCGTGGACACCCCGACCTGGGAAAGATGATCGCCCGATAAAGACGG	1980
QY	1919	GGGCTTGGGGGGCCCGTGGACACCCCGACCTGGGAAAGATGATCGCCCGATAAAGACGG	1978
QY	1981	ATTCTAAGG 1989	
QY	1979	ATTCTAAGG 1987	
RESULT 7			
AR204704		1987 bp	DNA
LOCUS	AR204704		linear
DEFINITION	Sequence 1 from patent US 6368796.		PAT 20-JUN-2002
ACCESSION	AR204704		
VERSION	AR204704.1	GI:21502099	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1987)		
AUTHORS	Avraham,H. and Groopman,J.E.		
TITLE	Methods of detection and treatment of breast cancer		
JOURNAL	Patent: US 6368796-A 1 09-APR-2002;		
FEATURES	location/qualifiers		
SOURCE	1..1987		
BASE COUNT	369 a 628 c 672 g 318 t		
ORIGIN			
Query Match	95.5%	Score 1909.8;	DB 6;
Best Local Similarity	99.2%	Pred. No. 9.5e-276;	Length 1987;
Matches 1973; Conservative	0;	Mismatches 7;	Indels 9;
		Gaps 5;	

[illegible]

```

Db      1086 ACCAGAGCGCGGCATGACAGACAGATGATCAACACAGAAACCTGGCGCTCTCTCTGCGGCTGA 1145
QY      1141 TCTTGACACGAGGGGCTGTATCATGTATGAGACAGCAGTGAAGGGGAAACCTGGTGAAT 1200
Db      1146 TCTTGACACGAGGGGCTGTATCATGTATGAGACAGCAGTGAAGGGGAAACCTGGTGAAT 1205
QY      1201 TTTCTGAGGACCGCGGGGCTGAGCCCTGCTGAACACCGCTGACCTCTGACATTTTCTCTGC 1260
Db      1206 TTTCTGAGGACCGCGGGGCTGAGCCCTGCTGAACACCGCTGACCTCTGACATTTTCTCTGC 1265
QY      1261 ACCTGGCCGAGGGGCTGAGTACTGAGAGGCAAGAGCTTGTGACACCGGACCTGGCCG 1320
Db      1266 ACCTGGCCGAGGGGCTGAGTACTGAGAGGCAAGAGCTTGTGACACCGGACCTGGCCG 1325
QY      1321 CCGGCAACATCTCTGCTCTCAGAGACCTGGTGGCCCAAGGTCACGACTTTGGGCTGGCCA 1380
Db      1326 CCGGCAACATCTCTGCTCTCAGAGACCTGGTGGCCCAAGGTCACGACTTTGGGCTGGCCA 1385
QY      1381 AACCCGAGCGGAAAGGGGCTAGACTCAAGCCGCTGCCCTCAAGTGAAGCGCCCGAGG 1440
Db      1386 AACCCGAGCGGAAAGGGGCTAGACTCAAGCCGCTGCCCTCAAGTGAAGCGCCCGAGG 1445
QY      1441 CTCTCAACGAGGGGAGTTCACCAAGATCGGATGTCTGAGATTTTGGGCTGCTCT 1500
Db      1446 CTCTCAACGAGG--GTTTACCAAGGAGTGGATGTCTGAGATTTTGGGCTGCTCT 1502
QY      1501 GGGAGGCTCTCTCATATGAGAGGGGCTCCGTAACCTTAAATGTCTCAAGAAAGAGTGGCG 1560
Db      1503 GGGAGGCTCTCTCATATGAGAGGGGCTCCGTAACCTTAAATGTCTCAAGAAAGAGTGGCG 1562
QY      1561 AGGCGGTGAGAAAGGGGTACCGCATGAAACCCCGGAGGGCTGTCCAGGCCCGTGCAGC 1620
Db      1563 AGGCGGTGAGAAAGGGGTACCGCATGAAACCCCGGAGGGCTGTCCAGGCCCGTGCAGC 1622
QY      1621 TCTCATGAGACACTGCTGGGAGGAGGAGCGCGCGCGCGGCAACCTTCCGAACTGG 1680
Db      1623 TCTCATGAGACACTGCTGGGAGGAGGAG--CGCGCGCGCGGCAACCTTCCGAACTGG 1681
QY      1681 CCGAGAGAGTGGCCCGGAGAGTACGAGAGTGCAGAGTCCCGGCTCCGTTCCAGGCGAG 1740
Db      1682 CCGAGAGAGTGGCCCGGAGAGTACGAGAGTGCAGAGTCCCGGCTCCGTTCCAGGCGAG 1741
QY      1741 AGCCGAGAGGCTCCACCTGCGCCCGAAGCAGAGCCCTGACCCCAACCGGGTGGGCGCT 1800
Db      1742 AGCCGAGAGG--TTCACCTGCGCCCGAAGCAGAGCCCTGACCCCAACCGCGCT--GGCCCT 1798
QY      1801 TGGCCCAAGAGGAGAGAGTGAAGTGGGCGGTGGGGGAGTGAACAGGCCCAAGG 1860
Db      1799 TGGCCCAAGAGGAGAGAGTGAAGTGGGCGGTGGGGGAGTGAACAGGCCCAAGG 1858
QY      1861 AGGCTCCAGGCGGAGAGTATCTCTCTGTTGCCACAGAGAGGGCTGGGCCACGTAAGG 1920
Db      1859 AGGCTCCAGGCGGAGAGTATCTCTCTGTTGCCACAGAGAGGGGCTGGGCCACGTAAGG 1918
QY      1921 GGGCTTGGGCGGCGGCTGAGACACCCAGACAGCTGCAAGAGATGATCCCGGATTAAGACG 1980
Db      1919 GGGCTTGGGCGGCGGCTGAGACACCCAGACAGCTGCAAGAGATGATCCCGGATTAAGACG 1978
QY      1981 ATTCTAAGG 1989
Db      1979 ATTCTAAGG 1987

```

```

RESULT 8
LOCUS      HUMMATK      1987 bp      mRNA      linear      PRI 11-MAR-1994
DEFINITION Human tyrosine kinase (MATK) mRNA, complete cds.
ACCESSION  U18974
VERSION     U18974.1 GI:455449
KEYWORDS   cytoplasmic protein; tyrosine kinase.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens

REFERENCE
1 (Bases 1 to 1987)
AUTHORS   Bennett, B.D., Cowley, S., Jiang, S., London, R., Deng, B., Grabarek, J.,
Groopman, J.E., Goeddel, D.V. and Avraham, H. and Avraham, H.
TITLE      Identification and characterization of a novel tyrosine kinase from
megakaryocytes
JOURNAL    J. Biol. Chem. 269 (2), 1068-1074 (1994)
MEDLINE    94117408
PUBMED     8288563
COMMENT    Original
FEATURES
source
location/Qualifiers
1..1987
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="CMK11-5"
/dev_stage="adult"
/tissue="liver"
/gene="MATK"
<1..63
/gene="MATK"
263..1846
/gene="MATK"
/codon_start=1
/product="tyrosine kinase"
/protein_id="AA16703.1"
/db_xref="GI:459550"
/translation="MARGSLYSWRAFHGCDNAEELPRVSPFLRAMHPPPYSAKMT
RRAPGTQCTITCEHTRPKPGELAFKRGVYVLTLECKNSWFRVNHNSGEGILAA
GALRDEALSDAPKLSLMPFHFKISGDAVOLOPDEGLFLYRSASHPGIVYLC
SEGRDVIHYRVLRDGLTIDEAFECNLMDVEHYSKDGAICTKLVAPKRNHGRS
AEEELARAGMLNLOHLTGAOLGEBEFVYLOGEYLGKAVAKNKICDYTAQAFIDE
TYMTKMOHBNLYRLGLTILHGLTYVMEHVSIGNVNLFRGRALVMTADLOPSTL
HVAEGHETLESKLVYRDIAARITVSEDLVAKVSFGLAKERKGLDSRLPYKTA
PEALKHGTFSKSDVSWFVGLWVEFVYGRAPYPKMLKESVSEVGEVMEPEEGCP
PVHVLMSCEWAEPPAGHPNSAMVPSGVAVOVPSPSGRTPVHLAPRGAULTP
PGGPWPQRTREVERSAWGH"
568
polya_site
/gene="MATK"
BASE COUNT      369 a      628 c      672 g      318 t
ORIGIN
Query Match      95.5%; Score 1909.8; DB 9; Length 1987;
Best Local Similarity 99.2%; Pred. No. 9.5e-276;
Matches 1973; Conservative 0; Mismatches 7; Indels 9; Gaps 5;
QY      1 CTGCTCCAAAGTTGTGCAAGCGGAGCCGCTCGGGGTGTGACGCGGCTGCGGAGGCC 60
Db      8 CTGCTCCAAAGTTGTGCAAGCGGAGCCGCTCGGGGTGTGACGCGGCTGCGGAGGCC 67
QY      61 TCTTGGGGGCGGGGCGGGGCGGCTGCGGGGCGGCGGCTGAGAGAAAGAAAGAAAGAAC 120
Db      68 TCTTGGGGGCGGGGCGGGGCGGCTGCGGGGCGGCGGCTGAGAGAAAGAAAGAAAGAAC 127
QY      121 AGGCTGGTCCAGTGGACACCCAGCTCTGCTGTGTCAGACCGGCTGGCTGGTGGCA 180
Db      128 AGGCTGGTCCAGTGGACACCCAGCTCTGCTGTGTCAGACCGGCTGGCTGGTGGCA 187
QY      181 GGGCATTTCCAGCGTCCCGAGCTGTGACCACTTGTCACTGTGCTTCACTGCTGCTGAG 240
Db      188 GGGCATTTCCAGCGTCCCGAGCTGTGACCACTTGTCACTGTGCTTCACTGCTGCTGAG 247
QY      241 TTTCCCTCTGGGGGCGGAGTGGCGGGGCGGAGGCTCTGTCTTCTCTGGGCGGATTTACAG 300
Db      248 TTTCC--TCTGTGGGGGAGTGGCGGGGCGGAGGCTCTGTCTTCTCTGGGCGGATTTACAG 305
QY      301 GGTGTGATTCGTGGTGAAGAACTTCCCGGGGTGAAGCCCGGCTTCTCCGAGGCTGGACCC 360
Db      306 GGTGTGATTCGTGGTGAAGAACTTCCCGGGGTGAAGCCCGGCTTCTCCGAGGCTGGACCC 365

```

QY 361 CCCCTCCGCTCTACAGCAGATGCCAAGAGGCGCTGGGCCCGGGGCAACCCAGTGTATCA 420
 Db 366 CCCCTCCGCTCTACAGCAGATGCCAAGAGGCGCTGGGCCCGGGGCAACCCAGTGTATCA 425
 QY 421 CCAAAATGGAGCACAACCCGCCCCAAGCAGGAGGAGCTGGCTTCCGCAAGGGGAGCTGAG 480
 Db 426 CCAAAATGGAGCACAACCCGCCCCAAGCAGGAGGAGCTGGCTTCCGCAAGGGGAGCTGAG 485
 QY 481 TCACATCTGGAGGCTGGCAGAACAGAGCTGTACCGCTGCAAGAACCAACACAGAGG 540
 Db 486 TCACATCTGGAGGCTGGCAGAACAGAGCTGTACCGCTGCAAGAACCAACACAGAGG 545
 QY 541 GACAGAGAGGAGCTGGTGGAGAGCTGGGAGGCTGGGAGAGGAGGAGGAGGAGGAGGAGG 600
 Db 546 GACAGAGAGGAGCTGGTGGAGAGCTGGGAGGCTGGGAGAGGAGGAGGAGGAGGAGGAGG 605
 QY 601 CCAAGCTCAGCTCATGCCGTGGTTCACAGGAGAGATCTGGGCCAAGAGAGCTGTCCAGC 660
 Db 606 CCAAGCTCAGCTCATGCCGTGGTTCACAGGAGAGATCTGGGCCAAGAGAGCTGTCCAGC 665
 QY 661 AGCTGAGGCTGGCAGAGATGGGCTGTTCTGTGGTGGGAGTCCGGCGGCGCACCCCGGCG 720
 Db 666 AGCTGAGGCTGGCAGAGATGGGCTGTTCTGTGGTGGGAGTCCGGCGGCGCACCCCGGCG 725
 QY 721 ACTAGCTCCTGTGCTGAGCTTTGGCCGAGCAGTCACTACCTACCGGCTGTGACACCGCG 780
 Db 726 ACTAGCTCCTGTGCTGAGCTTTGGCCGAGCAGTCACTACCGGCTGTGACACCGCG 785
 QY 781 ACGGCGACCTCAATTCATGAGGCGGTGTTCTTGCACCTCATGACATGTGTGAGC 840
 Db 786 ACGGCGACCTCAATTCATGAGGCGGTGTTCTTGCACCTCATGAGATGTGTGAGC 845
 QY 841 ATTACAGCAAGAGCAAGGGGCGCTATCTGACCAAGTGTGTGAGCAAAAGCGAAACAG 900
 Db 846 ATTACAGCAAGAGCAAGGGGCGCTATCTGACCAAGTGTGTGAGCAAAAGCGAAACAG 905
 QY 901 GGACCAAGTGGCGCAGAGAGAGTGGCGAGGCGGCGTGTACTGAACCTGACAGCATT 960
 Db 906 GGACCAAGTGGCGCAGAGAGAGTGGCGAGGCGGCGTGTACTGAACCTGACAGCATT 965
 QY 961 TGACATTGGGAGACACAGATCGAGAGGAGAGATTGGAGCTTCTTGCAGGGGTGAGTACC 1020
 Db 966 TGACATTGGGAGACACAGATCGAGAGGAGAGATTGGAGCTTCTTGCAGGGGTGAGTACC 1025
 QY 1021 TGGGGCAAAAGGTGGCGCGTGAAGATATGAAGTGTGATGTGACAGCCCGGCTTCTGCG 1080
 Db 1026 TGGGGCAAAAGGTGGCGCGTGAAGATATGAAGTGTGATGTGACAGCCCGGCTTCTGCG 1085
 QY 1081 ACGAGACGGCGCTCATGAGAGATGCAACACAGAAACCTGGTGGCTCTCTGAGCGTGA 1140
 Db 1086 ACGAGACGGCGCTCATGAGAGATGCAACACAGAAACCTGGTGGCTCTCTGAGCGTGA 1145
 QY 1141 TCCTCACCAGGGGCTGTATCTGTATGTGAGACAGTGTGAGCAAGGGCAACCTGTGTACT 1200
 Db 1146 TCCTCACCAGGGGCTGTATCTGTATGTGAGACAGTGTGAGCAAGGGCAACCTGTGTACT 1205
 QY 1201 TTCTGCGAGACCGGGGCTGAGACCTGTGTAACACCGCTAGCTCTCTGCGAGTTTCTCTGC 1260
 Db 1206 TTCTGCGAGACCGGGGCTGAGACCTGTGTAACACCGCTAGCTCTCTGCGAGTTTCTCTGC 1265
 QY 1261 ACGTGGCGGAGGGATGAGTACTGTGAGAGCAAGAAGCTTGTGCAACCGGAGCTGTGCGG 1320
 Db 1266 ACGTGGCGGAGGGATGAGTACTGTGAGAGCAAGAAGCTTGTGCAACCGGAGCTGTGCGG 1325
 QY 1321 CCGGGAACATCTGTGCTCAGAGAGCTGTGGCCAAAGGTTCAGCAGCTTTGGCTGTGCCA 1380
 Db 1326 CCGGGAACATCTGTGCTCAGAGAGCTGTGGCCAAAGGTTCAGCAGCTTTGGCTGTGCCA 1385
 QY 1381 AAGCCGAGCGAAGGGGCTAGACTCAAGCGGCTGCCCTCAAGTGTGAGCGCGCCGAGG 1440
 Db 1386 AAGCCGAGCGAAGGGGCTAGACTCAAGCGGCTGCCCTCAAGTGTGAGCGCGCCGAGG 1445
 QY 1441 CTCTCAAAACAGCGGAAGTTCAACAGCAAGTGTGATGTGTGAGGTTTGGGGTGTGCTCT 1500

Db 1446 CTCTCAAAACAGG---GTTCAACAGCAAGTGGATGTCTGGAGTTTGGGGTGTGCTCT 1502
 QY 1501 GGGAGCTTCTCATATATGAGAGCGGCTCCGTACCTTAAATGTCACTGAAAGAGTGTGG 1560
 Db 1503 GGGAGCTTCTCATATATGAGAGCGGCTCCGTACCTTAAATGTCACTGAAAGAGTGTGG 1562
 QY 1561 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGGGCTGTCCAGGCCCGGTGCAG 1620
 Db 1563 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGGGCTGTCCAGGCCCGGTGCAG 1622
 QY 1621 TCCTCATGAGCAGCTGTGGAGAGCAGAGCCCGCGCGGCGCACCTTCCGCAAACTGG 1680
 Db 1623 TCCTCATGAGCAGCTGTGGAGAGCAGAGCCCGCGCGGCGCACCTTCCGCAAACTGG 1681
 QY 1681 CCGAAGAGCTGGCCCGGAGCTACGCACTGTGCAAGTGTGCCCCAGCTTCCGTCTAGGGCAG 1740
 Db 1682 CCGAAGAGCTGGCCCGGAGCTACGCACTGTGCAAGTGTGCCCCAGCTTCCGTCTAGGGCAG 1741
 QY 1741 ACGCGGAGGCTTCACACTGCGCCCGGAGAGCAGAGCCTTGACCCGCGGAGGAGGAGGCT 1800
 Db 1742 ACGCGGAGGCTTCACACTGCGCCCGGAGAGCAGAGCCTTGACCCGCGGAGGAGGAGGCT 1798
 QY 1801 TGGCCCAAGAGAGCAGAGAGTGTGAGAGTGGCGGTGGGGGCACTGACAGAGCCCAAG 1860
 Db 1799 TGGCCCAAGAGAGCAGAGAGTGTGAGAGTGGCGGTGGGGGCACTGACAGAGCCCAAG 1858
 QY 1861 AGGCTTCAGAGCGGCGCAATCTCTCTGTGTGTGCTCCACAGAGAGGGCTGGCCAGTGG 1920
 Db 1859 AGGCTTCAGAGCGGCGGCAATCTCTCTGTGTGTGCTCCACAGAGAGGGCTGGCCAGTGG 1918
 QY 1921 GGCTGTGGGCGGCGCGGTGAGACACCCAGACCTGTGCAAGAGATGTGCGCGATTAAGACG 1980
 Db 1919 GGCTGTGGGCGGCGCGGTGAGACACCCAGACCTGTGCAAGAGATGTGCGCGATTAAGACG 1978
 QY 1981 ATCTTAAG 1989
 Db 1979 ATCTTAAG 1987

RESULT 9
 HSHYLRK
 LOCUS H. sapiens HYL tyrosine kinase mRNA. 1968 bp mRNA linear PRI 10-OCT-1994
 DEFINITION X77278
 ACCESSION X77278
 VERSION 1
 KEYWORDS HYLTK gene; nonreceptor protein tyrosine kinase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Sakano, S., Iwama, A., Inazawa, J., Ariyama, T., Ohno, M. and Suda, T.
 TITLE Molecular cloning of a novel non-receptor tyrosine kinase, HYL (hematopoietic consensus tyrosine-lacking kinase)
 JOURNAL Oncogene 9 (4), 1155-1161 (1994)
 MEDLINE 94181267
 PUBMED 8134117
 REFERENCE 2 (bases 1 to 1968)
 AUTHORS Iwama, A.
 DIRECT SUBMISSION Direct Submission
 JOURNAL Submitted (14-JAN-1994) A. Iwama, Dept of Cell Differentiation, Inst of Mol Embryology & Genetics, Kumamoto University School of Medicine, 2-2-1 Honjo, Kumamoto 860, JAPAN
 FEATURES
 location/Qualifiers
 1..1968
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="19p13"
 /tissue_type="hematopoietic cell"
 /clone_lib="Ur-7 cDNA library"
 1..1968
 gene

mrna
CDS
/gene="HHYLTk"
1..1968
/gene="HHYLTk"
208..1731
/gene="HHYLTk"
/codon.start=1
/product="Hyl tyrosine kinase"
/protein_10="CAA5493.1"
/db_xref="GI:557272"
/translation="MAGRSLSVSRFAHGDSDAELEPRVSPFLRAHPPVSAEMPT
RRAPGTCITKCEHTRPKPGCELAFRKGDVLTILEACENKSHRYVKNHTSGOSELILA
GALAEERALSNDPKLSMPFHGKISQEAVALQIDPEDEGLVRESARHGDVLY
SFGHDVTHIRLHEDGHLTIDEAFECNLMDWEHYSKIDGALCTKLVRKRKHTS
AEERLARGMLNLOHLTJGAQJEGEFAVLOGEYQKRVAVANINICDYTAQAFIDE
TAVMTKQHEMLVLLGLVILHOGLYIVMEHYSKSNLVNPLTRRALVNTAQLQFSL
HVAEMKYELESKLVHRLDLARNILVEDIVAKVSDGLKAEKKGDSRLPYKMTA
PEALKHGFSTKSDVMSFEGYLIMEVFEYGRAPYPMKSLKVESEVKEGMEPEGP
GPHVHLSSCWHEPAPRRPFRKIAEKLAELELRNAGAPASVSGDDADGSTSPNSQP"

BASE COUNT 391 a 614 c 652 g 311 t
ORIGIN

Query Match 95.5% Score 1909.4; DB 9; Length 1968;
Best Local Similarity 99.7% Pred. No. 11e-275;
Matches 1945; Conservative 0; Mismatches 1; Indels 5; Gaps 3;

52 CGGAGGCGCTCTCTGGGGGGGGCGGGG--GCGGCTCGGGGGCGCCCTCGAGCAGAAA 109
1 CGGAGGCGCTCTCTGGGGGGGGGGCGGGCGCGCTCGGGGGGGCGCCCTCGAGCAGAAA 60
110 CAGGAAGAACCGGCTGGGTCCAGTGGCAGCCAGCTCCCTACCTCTCTGTGCGAGCCCT 169
61 CAGGAAGAACCGGCTGGGTCCAGTGGCAGCCAGCTCCCTACCTCTCTGTGCGAGCCCT 120
170 GGCCTGTGGAGGCGCATTCAGGCGTCCCGACGTGTACCACTGTCTAGTGGCTCTC 229
121 GGCCTGTGGAGGCGCATTCAGGCGTCCCGACGTGTACCACTGTCTAGTGGCTCTC 180
230 ACCTGCTCAGTTTCTCTGTGGGGGGGAGTGGCGGGGAGGCTCTCTGTGTTTCTGGCG 289
181 ACCTGCTCAGTTTCTCTGTGGGGGGGAGTGGCGGGGAGGCTCTCTGTGTTTCTGGCG 239
290 GGCATTTCAGGCGGTATTCGTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCCG 349
240 GGCATTTCAGGCGGTATTCGTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCCG 299
350 AGCCTGGACCCCTCTCCGTCTCAGCAGAGATGCCAAGAGGGGCTGGGCCCGGCGAC 409
300 AGCCTGGACCCCTCTCCGTCTCAGCAGAGATGCCAAGAGGGGCTGGGCCCGGCGAC 359
410 CCAAGTATACCAAAATGCGAGCACACCCGCCCAAGCAGGGAGCTGSCCTTCCGAA 469
360 CCAAGTATACCAAAATGCGAGCACACCCGCCCAAGCAGGGAGCTGSCCTTCCGAA 419
470 GGGGAGCTGGTACCATCTCTGGAGGCTTGGAGAAAGAGCTGTACCGGCTCAAGCA 529
420 GGGGAGCTGGTACCATCTCTGGAGGCTTGGAGAAAGAGCTGTACCGGCTCAAGCA 479
530 CCAACACAGTGAAGAGAGGGGCTGTGAGCTGGGGCGCTGGGAGAGGGAGGCGCT 589
480 CCAACACAGTGAAGAGAGGGGCTGTGAGCTGGGGCGCTGGGAGAGGGAGGCGCT 539
590 CTCGCGAGACCCCAAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 649
540 CTCGCGAGACCCCAAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
650 GGGTTCACAGCTGAGCTCCCGAGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709
600 GGGTTCACAGCTGAGCTCCCGAGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
710 CCAACCCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769
660 CCAACCCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719

770 GCTGCACCGGAGCGCCACCTCACAATGATGAGCCGCTTCTTGTCAACCTCATGA 829
720 GCTGCACCGGAGCGCCACCTCACAATGATGAGCCGCTTCTTGTCAACCTCATGA 779
830 CATGTGTGAGCATTACGACGAAGGAGCAAGGCGCTATGTGCACCAAGCTGTGAGCCAA 889
780 CATGTGTGAGCATTACGACGAAGGAGCAAGGCGCTATGTGCACCAAGCTGTGAGCCAA 839
890 GCGGAACACGGGAGCAAGTGTGCGCGGAGGAGGCTGCGCAGGGCGGGGCTGATCGAA 949
840 GCGGAACACGGGAGCAAGTGTGCGCGGAGGAGGCTGCGCAGGGCGGGGCTGATCGAA 899
950 CTTGCACACATTGACATTGGGAGACAGATCGAGAGGAGAGTTTGGAGCTGTCTGCA 1009
900 CTTGCACACATTGACATTGGGAGACAGATCGAGAGGAGAGTTTGGAGCTGTCTGCA 959
1010 GGGTGTATCTCTGGGCGCAAAAGTGTGCGCGGAGAGATATCAAGTGTATGTGACGCCA 1069
960 GGGTGTATCTCTGGGCGCAAAAGTGTGCGCGGAGAGATATCAAGTGTATGTGACGCCA 1019
1070 GGCCTTCTGTGAGAGAGCGGCGCTCATGAGAAAGATGCAACAGAGAACTGGTGGCT 1129
1020 GGCCTTCTGTGAGAGAGCGGCGCTCATGAGAAAGATGCAACAGAGAACTGGTGGCT 1079
1130 CTTGGGCGGTATCTGTGACCAAGGCGCTGTACATTTGTATGAGACGTGAGCAAGGCAA 1189
1080 CTTGGGCGGTATCTGTGACCAAGGCGCTGTACATTTGTATGAGACGTGAGCAAGGCAA 1139
1190 CTTGTGTACATTTCTGTGCGGACCGGGGCTGCAAGCCCTCTGTACACCGCTTCTGCA 1249
1140 CTTGTGTACATTTCTGTGCGGACCGGGGCTGCAAGCCCTCTGTACACCGCTTCTGCA 1199
1250 GTTTTCTCTGACAGTGGCGGAGGAGGATGAGTACCTGAGAGAGCAAGAGCTTGTACACG 1309
1200 GTTTTCTCTGACAGTGGCGGAGGAGGATGAGTACCTGAGAGAGCAAGAGCTTGTACACG 1259
1310 CGACCTGGCGCGCGCAACATCTGTGTCTCAGAGAGACCTGTGGCCAAAGTACAGCACTT 1369
1260 CGACCTGGCGCGCGCAACATCTGTGTCTCAGAGAGACCTGTGGCCAAAGTACAGCACTT 1319
1370 TGGCGTGGCGCAAAAGCGGAGGAGGAGGCTGAGACTCAAGCGGCTGCGGCTCAAGTGGAC 1429
1320 TGGCGTGGCGCAAAAGCGGAGGAGGAGGCTGAGACTCAAGCGGCTGCGGCTCAAGTGGAC 1379
1430 GCGCGCGAGGCTCTCAAAACAGGGAAGTTCACACAGCAAGTTCGATGTGAGAGTTTGG 1489
1380 GCGCGCGAGGCTCTCAAAACAGGGAAGTTCACACAGCAAGTTCGATGTGAGAGTTTGG 1439
1490 GGTGTGTCTGTGGAAGCTTCTCATATGAGAGGCGCTCGTACCTTAATATGTCTGAA 1549
1440 GGTGTGTCTGTGGAAGCTTCTCATATGAGAGGCGCTCGTACCTTAATATGTCTGAA 1499
1550 AGAGGTGTGGAAGGCGTGTGAGAAAGGAGTTCGCAATGAGAAACCCCGGAGGCGTGTACAG 1609
1500 AGAGGTGTGGAAGGCGTGTGAGAAAGGAGTTCGCAATGAGAAACCCCGGAGGCGTGTACAG 1559
1610 CCGCGTGCAGCTCTCATGAGACGTGTGAGAGGAGAGAGCCCGCGCGCGCACCTT 1669
1560 CCGCGTGCAGCTCTCATGAGACGTGTGAGAGGAGAGAGCCCGCGCGCGCACCTT 1619
1670 CCGCAAACTGTGCGGAGAGTGGCGCGGAGCTACGCAAGTGTGAGTGTGCGCGCACCTT 1729
1620 CCGCAAACTGTGCGGAGAGTGGCGCGGAGCTACGCAAGTGTGAGTGTGCGCGCACCTT 1679
1730 CTTGAGGAGAGAGCGCGAGGCTGTCACTGTGCGCGCGCAAGAGAGCCCTGACCCACCC 1789
1680 CTTGAGGAGAGAGCGCGAGGCTGTCACTGTGCGCGCGCAAGAGAGCCCTGACCCACCC 1739
1790 GGTGGGCGCTTGGCGCGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1849
1740 GGT--GGCGCTTGGCGCGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1797

Accession	Sequence	Position
Oy	1850 CAGGCCCAAGAGAGGGTCCAGGCGGGCAAGTATCTCTCGTGGCCACAGCAGGAGGGCTGG	1909
Db	1798 CAGGCCCAAGAGAGGGTCCAGGCGGGCAAGTATCTCTCGTGGCCACAGCAGGAGGGCTGG	1857
Oy	1910 CCCACGTAGGAGGGGCTCTGGGCGGGCCGTGGACACCCAGACCTCGGAAGATGATCGCCC	1969
Db	1858 CCCACGTAGGAGGGGCTCTGGGCGGGCCGTGGACACCCAGACCTCGGAAGATGATCGCCC	1917
Oy	1970 GATAAAGCAGGATCTTAAGCACTTAATAAAA	2000
Db	1918 GATAAAGCAGGATCTTAAGCACTTAATAAAA	1948

	RESULT	10			
AR053292					
LOCUS					
DEFINITION	AR053292	1942 bp	DNA		
ACCESSION	Sequence	11 from patent US 5834208.			
VERSION	AR053292				
KEYWORDS	AR053292.1	GI:5978154			
SOURCE	.				
ORGANISM	Unknown.				
REFERENCE	Unknown.				
AUTHORS	Unclassified.				
TITLE	1 (bases 1 to 1942)				
JOURNAL	Sakano,S.				
FEATURES	Tyrosine kinase				
Source	Patent: US 5834208-A	11 10-NOV-1998;			
	Location/Qualifiers				
	1..1942				
	/organism="unknown"				
BASE COUNT	365 a	614 c	652 g	311 t	
ORIGIN					

Query Match	95.2%	Score 1903.4	DB.6	Length 1942
Best Local Similarity	99.7%	Pred. No. 8.6e-275		
Matches 1039, Conservative	0	Mismatches 1	Indels 5	Gaps 3

Oy	52	GGAGAGCCCTCTTGGGGGCGGGGCGGG - GGGGCTGGGGGGGCGCCCTTGAGAGAAA	109
Dd	1	CGAGAGCCCTCTTGGGGGCGGGGCGGGGCGGGCTCGGGGGGCGCCCTTGAGAGAAA	60
Oy	110	CAGGAAAGAACAGGCTCGTGGTCCAGTGGCACCCAGCTCCCTACTCTGTGGCAGCGGCT	169
Dd	61	CAGGAAAGAACAGGCTCGTGGTCCAGTGGCACCCAGCTCCCTACTCTGTGGCAGCGGCT	120
Oy	170	GGCGTGGGAGGCGCATTCGAGGGTCCCGGACGTGAGACCACTTGACAGTGTGCTTC	229
Dd	121	GGCGTGGGAGGCGCATTCGAGGGTCCCGGACGTGAGACCACTTGACAGTGTGCTTC	180
Oy	230	ACCTGGCTCAGTTTCCCTCTGGGGGGCGATGGCGGGGGGAGGCTCTGTGTTTCTTGCG	289
Dd	181	ACCTGGCTCAGTTTCCCTCT - GGGGGGAGTGGCGGGGGGAGGCTCTGTGTTTCTTGCG	239
Oy	290	GGCATTTGACGGCTGTGATTTCTGCTGAGGAACCTTCCGGGTTGAGCCCGCTTCTCG	349
Dd	240	GGCATTTGACGGCTGTGATTTCTGCTGAGGAACCTTCCGGGTTGAGCCCGCTTCTCG	299
Oy	350	AGCGTGGACCCCTCCCTCTCTAGCGCAGGATGCAACGAGGGGCTGGGCCCGGGGAC	409
Dd	300	AGCGTGGACCCCTCCCTCTCTAGCGCAGGATGCAACGAGGGGCTGGGCCCGGGGAC	359
Oy	410	CCAAGTGTATCACCAATGCGAGCACACC CGCCCAAGCAGGGGAGCTGGCTTCCGCAA	469
Dd	360	CCAAGTGTATCACCAATGCGAGCACACC CGCCCAAGCAGGGGAGCTTCCGCAA	419
Oy	470	GGGGAGACTGGGACACATCTGTGGAGGCTTGGGAGAACAGAGTGGTATCCGCTCAACGA	529
Dd	420	GGGGAGACTGGGACACATCTGTGGAGGCTTGGGAGAACAGAGTGGTATCCGCTCAACGA	479
Oy	530	CCACACCAAGTGGACAGAGGGGGCTGGTGGCACTGGGGGCTGGCGGAGCGGGAGCCCT	589
Dd	480	CCACACCAAGTGGACAGAGGGGGCTGGTGGCACTGGGGGCTGGCGGAGCGGGAGCCCT	539

QY	530	CTCCGACAGACCCCAAGCTCAAGCTCATGTCGGTGTTCACAGGGAAGATTCTCGGGCCACGA	649
Db	540	CTCCGACAGACCCCAAGGCTCAAGCTCATGTCGGTGTTCACAGGGAAGATTCTCGGGCCACGA	599
QY	650	GGCGTCCAGAGGTGGAGCCCTCCCGAAGATGGGCTGTTCCTGGATGGGGAGTCCCGGCG	709
Db	600	GGCGTCCAGAGGTGGAGCCCTCCCGAAGATGGGCTGTTCCTGGATGGGGAGTCCCGGCG	659
QY	710	CCACCCGGGACATACGCTCTGTGCGAGCTTTGGCGGGACGCTCATCCACTACCGCT	769
Db	660	CCACCCGGGACATACGCTCTGTGCGAGCTTTGGCGGGACGCTCATCCACTACCGCGT	719
QY	770	GCTCAGCCGAGCGGACCTCACATCGATGAGCCGTGTTCTTGTGCAACCTCATGGA	829
Db	720	GCTCAGCCGAGCGGACCTCACATCGATGAGCCGTGTTCTTGTGCAACCTCATGGA	779
QY	830	CATGTGGAGCATTAACAGCAAGGAGGCGCTATCTGCACCAAGCTGGTGAAGCCAAA	889
Db	780	CATGTGGAGCATTAACAGCAAGGAGGCGCTATCTGCACCAAGCTGGTGAAGCCAAA	839
QY	890	GCAGAAACACGGGACCAAGTGGCGCGAGAGAGACTGGCCAGGGCGGGCTGTACTGAA	949
Db	840	GCAGAAACACGGGACCAAGTGGCGCGAGAGAGACTGGCCAGGGCGGGCTGTACTGAA	899
QY	950	CCTCAGACATTTGACATTTGGGAGCACAAGATCGAGAGAGAGATTTGGAGCTGTCTGCA	1009
Db	900	CCTCAGACATTTGACATTTGGGAGCACAAGATCGAGAGAGAGATTTGGAGCTGTCTGCA	959
QY	1010	GGGTGAGTACTGGGGGCAAAAGGTGGCGCGTGAAGATATCAAGTGTGATGTGACAGCCA	1069
Db	960	GGGTGAGTACTGGGGGCAAAAGGTGGCGCGTGAAGATATCAAGTGTGATGTGACAGCCA	1019
QY	1070	GGCCCTCTTGACAGAGAGCGCCGTATGACCAAGTCAACACAGAAACCTGGGCGCTCT	1129
Db	1020	GGCCCTCTTGACAGAGAGCGCCGTATGACCAAGTCAACACAGAAACCTGGGCGCTCT	1079
QY	1130	CCTGGGCGTGATCTCTGCACCAAGGGGCTGTACATTTGTATGGAGCAGCTGAGCAAGGCCAA	1189
Db	1080	CCTGGGCGTGATCTCTGCACCAAGGGGCTGTACATTTGTATGGAGCAGCTGAGCAAGGCCAA	1139
QY	1190	CCTGGTGAACTTTCTGGCGGAGCCGGGGTCCAGCCCTGCTGAACAACCGCTCAGCTCCTGCA	1249
Db	1140	CCTGGTGAACTTTCTGGCGGAGCCGGGGTCCAGCCCTGCTGAACAACCGCTCAGCTCCTGCA	1199
QY	1250	GTTTTCCTTGACGCGGGCCGAGGGCATGAGATCCTGGAAGCAAGAAGCTGTGCAACCG	1309
Db	1200	GTTTTCCTTGACGCGGGCCGAGGGCATGAGATCCTGGAAGCAAGAAGCTGTGCAACCG	1259
QY	1310	CGACCTGGCCGCCCGCAACATCTGTGCTCAGAGAGACCTGGTGGCCAAAGTCAAGCACTT	1369
Db	1260	CGACCTGGCCGCCCGCAACATCTGTGCTCAGAGAGACCTGGTGGCCAAAGTCAAGCACTT	1319
QY	1370	TGGCTTGCGCCAAAGCCGAGCGGAAGGGGCTTGACTCAAGCCGCTGCCCGTCAAGTGGAC	1429
Db	1320	TGGCTTGCGCCAAAGCCGAGCGGAAGGGGCTTGACTCAAGCCGCTGCCCGTCAAGTGGAC	1379
QY	1430	GGCGCCGAGGCTCTCAACACAGGGAAGTTACACGACAAGTCGATGCTGGAATTTGG	1489
Db	1380	GGCGCCGAGGCTCTCAACACAGGGAAGTTACACGACAAGTCGATGCTGGAATTTGG	1439
QY	1490	GGTGTGCTGTGGGAGGTCTTCTCATATGACGCGGCTCCGTACCCCTAAATGTCACTGAA	1549
Db	1440	GGTGTGCTGTGGGAGGTCTTCTCATATGAGAGGGGCTCCGTACCCCTAAATGTCACTGAA	1499
QY	1550	AGAGGTGTCGAGGCGCTGAGAGAGGGGTACCCGATGAAACCCCGAGGGCTGTCCAGG	1609
Db	1500	AGAGGTGTCGAGGCGCTGAGAGAGGGGTACCCGATGAAACCCCGAGGGCTGTCCAGG	1559
QY	1610	CCCCGTGACGTCCTCATGAGACACTGCTGGGAGGCAAGAGCCCGCCCGCGGCAACCTT	1669
Db	1560	CCCCGTGACGTCCTCATGAGACACTGCTGGGAGGCAAGAGCCCGCCCGCGGCAACCTT	1619
QY	1670	CCGGAACCTGGCCAGAAAGCTGGCGCCGGGAGCTACGAGTGCAGGTGCCAGTCCCGACGCTCCGT	1729

Db	1620	CCGCAAACTGGCCCGAAGCTGTGGCCCGGAGACTACAGTGCAGGTGGCCAGGCTCCGT	1679
Qy	1730	CTCAGAGCGAGACGCCCAACGGCTCACCCTCGCCCGAAGCCAGAGACCCCTGACCCACC	1789
Db	1660	CTCAGAGGCGAGACCGCCGACGGCTCCACTCCCTCCCGGAAGCCAGAGGCTTACCCCCACC	1739
Qy	1790	GGTGGGGGCTCTTGGCCCGACAGAGACCGAGAGTGGAGAGTGGCGGCTGGGGGGCAGTAC	1849
Db	1740	GGT--GGCCCTTGGCCCGACAGAGACCGAGAGTGGAGAGTGGCGGCTGGGGGGCAGTAC	1797
Qy	1850	CAGGCCCAAGAGGGGTCCAGGCGGGGCAAGTATCTCTCGTGGCCACAGCAGGAGGCTGG	1909
Db	1798	CAGGCCCAAGAGGGGTCCAGGCGGGGCAAGTATCTCTCGTGGCCACAGCAGGAGGCTGG	1857
Qy	1910	CCCAAGTATGGGGGCTCTGGGCGGGCCCGTGGACACCCAGACCTCGAAGATGATCGCC	1969
Db	1858	CCCAAGTATGGGGGCTCTGGGCGGGCCCGTGGACACCCAGACCTCGAAGATGATCGCC	1917
Qy	1970	GATTAAGACGATTTCTAAGGACTCT	1994
Db	1918	GATTAAGACGATTTCTAAGGACTCT	1942

FEATURES	source
LOCUS	E10345
DEFINITION	E10345 cDNA encoding human novel tyrosine kinase.
VERSION	E10345
KEYWORDS	E10345.1 GI:22027175
SOURCE	JP 1995313157-A/1
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 1942)
JOURNAL	Sakano S.
COMMENT	NEW TYROSINE KINASE
	Patent: JP 1995313157-A 1 05-DEC-1995;
	ASahi Chem Ind Co Ltd
	OS Homo sapiens (human)
	PN JP 1995313157-A/1
	PD 05-DEC-1995
	PF 11-AUG-1994 JP 1994189444
	PR 25-AUG-1993 JP 93P 210403, 29-MAR-1994 JP 94P 58553 PI
	SAKANO SEIJI
	PC C12N9/12,A61K38/45,C07K14/47,C12N15/09,(C12N9/12,C12R1:91), PC
	(C12N15/09,
	PC C12R1:91);
	CC strandedness: Double;
	CC topology: Linear;
	CC hypothetical: No;
	CC anti-sense: No;
	PH Key
	Location/Qualifiers
FT	source
FT	1..1942
FT	/organism='Homo sapiens'
FT	/cell_type='megakaryoblast'
FT	/cell_line='U97'
FT	5'UTR
FT	1..207
FT	CDS
FT	208..1731
FT	/product='human novel tyrosine kinase' FT
FT	1732..1942
FT	CDS
FT	331..1731
FT	/product='human novel tyrosine kinase' FT
FT	misc_RNA
FT	349..540
FT	/product='SH3 domain'
FT	misc_RNA
FT	571..795
FT	/product='SH2 domain'
FT	misc_RNA
FT	904..1641
FT	/product='tyrosine kinase domain'
	Location/Qualifiers
	1..1942

[illegible]


```

OY 1010 GGGTGAGTACCTGGGGCAAAAGTGGCCGTGAGAAATATCAAGTGTGATGTGACGCCCA 1069
    |||||||
DB 960 GGGTGTAGTACCTGGGGCAAAAGTGGCCGTGAGAAATATCAAGTGTGATGTGACGCCCA 1019
OY 1070 GGGCTTCTGTGACGAGAGCGCCGTCTATGACGAAGATGACACGAGAACCTGTGTGCTCT 1129
    |||||||
DB 1020 GGGCTTCTGTGACGAGAGCGCCGTCTATGACGAAGATGACACGAGAACCTGTGTGCTCT 1079
OY 1130 CCTGGGGCTGTATCTGTGACACGAGGGCTGTACTATTGTCTATGAGACAGCTGTGACGACAA 1189
    |||||||
DB 1080 CCTGGGGCTGTATCTGTGACACGAGGGCTGTACTATTGTCTATGAGACAGCTGTGACGACAA 1139
OY 1190 CCTGTGAACTTTCTGGGAGACCCGGGGTGTGAGCCCTGTGAAACACCGCTGTGCTGTGA 1249
    |||||||
DB 1140 CCTGTGAACTTTCTGGGAGACCCGGGGTGTGAGCCCTGTGAAACACCGCTGTGCTGTGA 1199
OY 1250 GTTTTCTGTGACGAGTGGCCGAGGGGATGAGTACTGTGAGAGACAGAAAGCTTGTGACCG 1309
    |||||||
DB 1200 GTTTTCTGTGACGAGTGGCCGAGGGGATGAGTACTGTGAGAGACAGAAAGCTTGTGACCG 1259
OY 1310 CGACTGTGCGCCCGCCCAACATCTGTGTCTGAGAGACCTGTGTGCGCAAGGTGACGACTT 1369
    |||||||
DB 1260 CGACTGTGCGCCCGCCCAACATCTGTGTCTGAGAGACCTGTGTGCGCAAGGTGACGACTT 1319
OY 1370 TGGCGTGGCCAAAGCGGAGCGGAGGGGCTGAGTCAAGCCGGGCTGCGGCTCAAGTGGAC 1429
    |||||||
DB 1320 TGGCGTGGCCAAAGCGGAGCGGAGGGGCTGAGTCAAGCCGGGCTGCGGCTCAAGTGGAC 1379
OY 1430 GGGCGCCGAGGCTCTCAAAACACGSGGAGTTCACGACGAAGTGGATGTCTGTGAGTTTGG 1489
    |||||||
DB 1380 GGGCGCCGAGGCTCTCAAAACACGSGGAGTTCACGACGAAGTGGATGTCTGTGAGTTTGG 1439
OY 1490 GGTGTGTCTGTGGGAGGTCTTCTCATATGAGAGGGCTCGGTACCTTAAATGTCACTGAA 1549
    |||||||
DB 1440 GGTGTGTCTGTGGGAGGTCTTCTCATATGAGAGGGCTCGGTACCTTAAATGTCACTGAA 1499
OY 1550 AGAGTGTGCGGAGCGGTGGAGAGGGGATACCGATGGAACCCCGGAGGGGCTGTCCAGG 1609
    |||||||
DB 1500 AGAGTGTGCGGAGCGGTGGAGAGGGGATACCGATGGAACCCCGGAGGGGCTGTCCAGG 1559
OY 1610 CCCCGTGCACGTCTCATATGAGAGCTGTGTGAGGACAGAGCCCGCCGCGCACCTT 1669
    |||||||
DB 1560 CCCCGTGCACGTCTCATATGAGAGCTGTGTGAGGACAGAGCCCGCCGCGCACCTT 1619
OY 1670 CCGCAAACTGGCCGAGAAAGCTGGCCGGGAGCTACGCAAGTGTGCCCCAGCCTCGT 1729
    |||||||
DB 1620 CCGCAAACTGGCCGAGAAAGCTGGCCGGGAGCTACGCAAGTGTGCCCCAGCCTCGT 1679
OY 1730 CTCAGGGGAGAGGCGGAGCGGCTGCACCTCGCCCGGAGGCGGAGGCGCTGACCCACCC 1789
    |||||||
DB 1680 CTCAGGGGAGAGGCGGAGCGGCTGCACCTCGCCCGGAGGCGGAGGCGCTGACCCACCC 1739
OY 1790 GGTGGGGCCCTTGGCCCGCAGAGGACCGAGAGTGTGAGAGTGTGGGGGCGACTGAC 1849
    |||||||
DB 1740 GGT--GGCCCTTGGCCCGCAGAGGACCGAGAGTGTGAGAGTGTGGGGGCGACTGAC 1797
OY 1850 CAGGCCCAAGAGGGTCTCAGGCGGCGCAAGTCACTCTCTGTGTGCCACAGAGGAGCTGG 1909
    |||||||
DB 1798 CAGGCCCAAGAGGGTCTCAGGCGGCGCAAGTCACTCTCTGTGTGCCACAGAGGAGCTGG 1857
OY 1910 CCCAGTGGGGGGCTGTGGGGGGCGCGTGGAGACACCCGAGACTGTGGGAAGGATGTGCGCC 1969
    |||||||
DB 1858 CCCAGTGGGGGGCTGTGGGGGGCGCGTGGAGACACCCGAGACTGTGGGAAGGATGTGCGCC 1917
OY 1970 GATTAAGACGATTTCTAAGGACTCT 1994
    |||||||
DB 1918 GATTAAGACGATTTCTAAGGACTCT 1942

```

RESULT 12
HSM802388
LOCUS

HSM802388

1804 bp

mRNA

linear

PRI 18-FEB-2000

```

DEFINITION Homo sapiens mRNA; cDNA DKFZp434N1212 (from clone DKFZp434N1212);
partial cds.
ACCESSION AL137754
VERSION AL137754.1 GI:6808456
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
  1 (bases 1 to 1804)
  Duesterhoeft,A., Lauber,J., Mewes,H.W., Well,B. and Wiemann,S.
  Direct Submission
  Submitted (15-JAN-2000) MIPS, Am Klopferstr. 18a, D-82152
  Martinsried, GERMANY
COMMENT
  Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
  Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
  sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
  consortium of the German Genome Project.
  This clone (DKFZp434N1212) is available at the RZPD in Berlin.
  Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
  Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
  information about the clone and the sequencing project is available
  at http://www.mips.biochem.mpg.de/proj/cDNA/.
  Location/Qualifiers
    1..1804
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /map="19p13"
      /clone="DKFZp434N1212"
      /tissue_type="testis"
      /clone_id="434 (synonym: htes3). Vector pSPORT1; host
      DH10B; sites NotI + SalI"
      /dev_stage="adult"
      1..1804
      /gene="DKFZp434N1212"
      <1..1574
      /note="DKFZp434N1212"
      /product="protein-tyrosine kinase (EC 2.7.1.112) (Homo
      sapiens)"
      /codon_start=3
      /product="hypotheetical protein"
      /protein_id="CAB70906.1"
      /db_xref="GI:6808457"
      /translation="GNGATRSKLSRDLGVCSSRAEILLNAGRGSGAPPEVOYSPR
      FLRAWHPVSAARMPTRRMARFQTCITCEHTRPPGSLARFGDVVITLACEKSM
      YRVKHTSGOGLLAAGLRRELSADPKLSLMPWFKLSGDAVOQLPDEGLF
      LVRESARHPGQVYLCSFGRDIVHYRVLRDGHLLIDEAVFCNLMDVENVSKDGA
      ICTKLVPRKRKHGTSASEEELARAGMLNLNHLTGADVIGSEFCNALVLOGEYLGKVA
      VKNKCDVYTAQAFIDETA.VMTKMOHEIN.VRLIGVLIHOGIIVYMHVSKNLVNFLEF
      RGRALVNTAQLDLSLHVAEGMEYLESKLVLRDLAARNILYSEDVAVKVDGLAKA
      ERKGLDSSRLPLVKWAPALHKGFTSSLDVASFVLLMEVFTGRAPVPMKSLKES
      EAVENGYSMEPEEGCPGPVHVLMSSCWAEAPRRPFRKLEKLRLRSAGAPASVS
      GODADGSTSPRSQEP"
      1782
      polyA_site
      /gene="DKFZp434N1212"
      BASE COUNT 364 a 557 c 614 g 269 t
      ORIGIN
      Query Match 83.5%; Score 1669.2; DB 9; Length 1804;
      Best Local Similarity 98.4%; Pred. No. 8.3e-240;
      Matches 1686; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
OY 287 GCGGCATTTTCAGGCGCTGTGATTTCTGTGAGAGACTTCCCGGGTGTAGCCCGCTTCT 346
    |||||
DB 80 GGGGGCGGGCGCGCGCGCTCGGAGCGCCCGCTGAGCGAGGAGCGCTTCT 139
OY 347 CCGAGCTGGGACCGCCCGCTCCGCTGACGACGATGCGCAACGAGGCGGTGGCGCCGGG 406
    |||||||
DB 140 CCGAGCTGGGACCGCCCGCTCCGCTGACGACGATGCGCAACGAGGCGGTGGCGCCGGG 199
OY 407 CACCGAGTGTATCACCAATGCGAGACACCGCCCGCAAGCCAGGCGGAGCTGGCTTCG 466
    |||||||

```

Db 200 CACCAAGTATACCAAAATGAGACACACCCGCCCAAGCCAGGGAGAGTGGCTTCCG 259
 Qy 467 CAAGGGGACGTGTGTACCATCTCTGGAGGCTTGGGAGAAACAAGCTGTATCCGCTCAA 526
 Db 260 CAAGGGGACGTGTGTACCATCTCTGGAGGCTTGGGAGAAACAAGCTGTATCCGCTCAA 319
 Qy 527 GCACCAACACAGTGTGAGAGAGGGGCTGTGAGAGTGGGGGCTGGCGGAGCCGGAGGC 586
 Db 320 GCACCAACACAGTGTGAGAGAGGGGCTGTGAGAGTGGGGGCTGGCGGAGCCGGAGGC 379
 Qy 587 CCTCTCGGACAGCCCAAGCTCAGCTTCATCCGCTGTGTCCAGCGGGAAGATCTCGGCGCA 646
 Db 380 CCTCTCGGACAGCCCAAGCTCAGCTTCATCCGCTGTGTCCAGCGGGAAGATCTCGGCGCA 439
 Qy 647 GAGAGCTGTCCAGAGCTGCAGCTCCCGAGAGATGGGCTGTCTCTGTGGGAGTCCGC 706
 Db 440 GAGAGCTGTCCAGAGCTGCAGCTCCCGAGAGATGGGCTGTCTCTGTGGGAGTCCGC 499
 Qy 707 GCGCACCCCGGAGCTACGTCTGTGGTGTAGGCTTGGCGGAGCGTCAATCCATCCG 766
 Db 500 GCGCACCCCGGAGCTACGTCTGTGGTGTAGGCTTGGCGGAGCGTCAATCCATCCG 559
 Qy 767 CGTGTCTCAGCCGCGACGCGCACTCAGATTCGATGAGGCGCTTCTTCTTCTCAACCTCAT 826
 Db 560 CGTGTCTCAGCCGCGACGCGCACTCAGATTCGATGAGGCGCTTCTTCTTCTCAACCTCAT 619
 Qy 827 GGAATATGTGAGGATTTACAGCAAGAGCAAGGCGCTATCTGACCAACAGCTGTGAGACC 886
 Db 620 GGAATATGTGAGGATTTACAGCAAGAGCAAGGCGCTATCTGACCAACAGCTGTGAGACC 679
 Qy 887 AAACGGGAAACACGGGAGCAAGTGGCGGAGAGAGAGTGGCGGCGGCTGTACT 946
 Db 680 AAACGGGAAACACGGGAGCAAGTGGCGGAGAGAGAGTGGCGGCGGCTGTACT 739
 Qy 947 GAACCTGCAGCATTGTGACATTGGAGACAGATGAGAGAGAGTGGAGCTGTCT 1006
 Db 740 GAACCTGCAGCATTGTGACATTGGAGACAGATGAGAGAGAGTGGAGCTGTCT 799
 Qy 1007 GAGAGGTGATGACTGGGGGAAAAAGTGGCGCTGTGAATAATATCAAGTGTATGTGACAGC 1066
 Db 800 GAGAGGTGATGACTGGGGGAAAAAGTGGCGCTGTGAATAATATCAAGTGTATGTGACAGC 859
 Qy 1067 CGAGGCTTCTCTGAGACGAGAGGCGCTGTATGACGAAATGCAACACAGAGAACTGGTGGC 1126
 Db 860 CGAGGCTTCTCTGAGACGAGAGGCGCTGTATGACGAAATGCAACACAGAGAACTGGTGGC 919
 Qy 1127 TCTCTCTGAGCTGTATCTGTGACCAAGGGGCTGTATGATGAGAGCACTGAGCAAGG 1186
 Db 920 TCTCTCTGAGCTGTATCTGTGACCAAGGGGCTGTATGATGAGAGCACTGAGCAAGG 979
 Qy 1187 CAACCTGTGAACTTCTGTGGGAAACCGGGGTCGAGCCCTGTGAACACCCCTCAAGTCTCT 1246
 Db 980 CAACCTGTGAACTTCTGTGGGAAACCGGGGTCGAGCCCTGTGAACACCCCTCAAGTCTCT 1039
 Qy 1247 GCAATTTCTCTGACAGTGGCGGAGGAGTGTATGAGAGTGTGAGAGCAAGAGCTGTGCA 1306
 Db 1040 GCAATTTCTCTGACAGTGGCGGAGGAGTGTATGAGAGTGTGAGAGCAAGAGCTGTGCA 1099
 Qy 1307 CCGGAGACTGGCGCCCGCAACATCTGTGTCTAGAGAGACTGGTGGCCAAAGTCAAGCA 1366
 Db 1100 CCGGAGACTGGCGCCCGCAACATCTGTGTCTAGAGAGACTGGTGGCCAAAGTCAAGCA 1159
 Qy 1367 CTTTGGGCTGTGCAAAAGCCGAGGAGGAGGCTGTAGACTCAAGCCGCTGCCCTCAAGTG 1426
 Db 1160 CTTTGGGCTGTGCAAAAGCCGAGGAGGAGGCTGTAGACTCAAGCCGCTGCCCTCAAGTG 1219
 Qy 1427 GAGGCGCCCGGAGGCTGTCAAAACAGGGAGTTCAACAGCAAGTGGAGTGTGAGGTTT 1486
 Db 1220 GAGGCGCCCGGAGGCTGTCAAAACAGGGAGTGTCAACAGCAAGTGGAGTGTGAGGTTT 1279
 Qy 1487 TGGGGTGTCTGTGGAGGCTTCTCATATGAGAGGCTGCCCTTACCTTAAATGTCTACT 1546
 Db 1280 TGGGGTGTCTGTGGAGGCTTCTCATATGAGAGGCTGCCCTTACCTTAAATGTCTACT 1339

Qy 1547 GAAGAGGTGTGAGAGGCGCTGAGAAAGGGTACCGCATGTGAAACCCCGAGGCTGTCTC 1606
 Db 1340 GAAGAGGTGTGAGAGGCGCTGAGAAAGGGTACCGCATGTGAAACCCCGAGGCTGTCTC 1399
 Qy 1607 AGGCGCCGTGTGACGTCTCTCATATGACAGCTGTGTGGAGGACAGACCCCGCCGCGCAC 1666
 Db 1400 AGGCGCCGTGTGACGTCTCTCATATGACAGCTGTGTGGAGGACAGACCCCGCCGCGCAC 1459
 Qy 1667 CTTCCGCAAACTGGCCGAGAGTGTGGCCCGGAGCTACGCGAGTGTGAGTGGCCAGGCTC 1726
 Db 1460 CTTCCGCAAACTGGCCGAGAGTGTGGCCCGGAGCTACGCGAGTGTGAGTGGCCAGGCTC 1519
 Qy 1727 CGTCTCAGGCGAGAGCGCGGCTCCACCTTCGCCCCGGAAGCCAGAGAGCCCTGACCCCA 1786
 Db 1520 CGTCTCAGGCGAGAGCGCGGCTCCACCTTCGCCCCGGAAGCCAGAGAGCCCTGACCCCA 1579
 Qy 1787 CCGGCTGGGGCCCTTGGCCCCCAGAGACCCGAGAGAGTGTGAGTGTGGCGGCTGTGAG 1846
 Db 1580 CCGGCTGGGGCCCTTGGCCCCCAGAGACCCGAGAGAGTGTGAGTGTGGCGGCTGTGAG 1639
 Qy 1847 GACCAGGCCCCAGAGAGGAGGTCAGGCGGAGCAAGTATCTCTGTGGCCACAGAGGAGG 1906
 Db 1640 GACCAGGCCCCAGAGAGGAGGTCAGGCGGAGCAAGTATCTCTGTGGCCACAGAGGAGG 1639
 Qy 1907 TGGCCACAGTAAAGGAGGCTGTGGGCGGCGCTGTGACACCCAGAGCTGTGAGATG 1966
 Db 1700 TGGCCACAGTAAAGGAGGCTGTGGGCGGCGCTGTGACACCCAGAGCTGTGAGATG 1759
 Qy 1967 CCGGATAAAGCGGATTTCTAAGGACTTAAAGAA 2000
 Db 1760 CCGGATAAAGCGGATTTCTAAGGACAAAGAAAAA 1793

RESULT 13
 AK055395
 LOCUS
 DEFINITION
 Homo sapiens cDNA FLJ30833 fis, clone FEBR200207, highly similar to MEGAKARYOCYTE-ASSOCIATED TYROSINE-PROTEIN KINASE (EC 2.7.1.112).
 AK055395
 VERSION
 AK055395.1 GI:16550112
 KEYWORDS
 oligo capping; fis (full insert sequence).
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens

REFERENCE
 AUTHORS
 1
 Kakamori, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Makamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagaitsuna, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 TITLE
 Unpublished
 JOURNAL
 Isogai, T., Otsuki, T. and Sugiyama, T.
 REFERENCE
 2 (bases 1 to 2073)
 AUTHORS
 Direct Submission
 JOURNAL
 Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Katsarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
 source
 location/Qualifiers
 1..2073
 /organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FEBRA200207"
/tissue_type="brain"
/tissue="FEBRA2"
/dev_stage="fetus"
/clone_id="FEBRA2"
/notes="cloning vector: pME18SFL3"
BASE COUNT      415 a      634 c      703 g      321 t
ORIGIN

```

```

Query Match      83.2% Score 1663; DB 9; Length 2073;
Best Local Similarity 99.4%; Pred. No. 6,7e-239;
Matches 1669; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 311 TCGTGAAGAACTTCCCGGGGTGAGCCCGGCTTCTCCGAGGCTGGACCCCGCTCCGCT 370
DB 395 TGGTACAGCAGCACTTCTGTGAGCCCGGCTTCTCCGAGGCTGGACCCCGCTCCGCT 454
QY 371 CTCAGCCAGATGCGCAACGAGCGCTGGGCCCCGGGACCCAGTGTATCACCAGATGCGA 430
DB 455 CTCAGCCAGATGCGCAACGAGCGCTGGGCCCCGGGACCCAGTGTATCACCAGATGCGA 514
QY 431 GCACACCCCGCCCAAGCCAGGAGCTGGCTTCCGCAAGGGCGAGCTGTACCATCTCT 490
DB 515 GCACACCCCGCCCAAGCCAGGAGCTGGCTTCCGCAAGGGCGAGCTGTACCATCTCT 574
QY 491 GGAGGCGCTGCGAAGCAAGAGCTGTACCGCGTCAAGCAGCAGCAGTGTGAGAGAGGG 550
DB 575 GGAGGCGCTGCGAAGCAAGAGCTGTACCGCGTCAAGCAGCAGCAGTGTGAGAGAGGG 634
QY 551 GCTGTGAGAGTGGGGCGCTCGGAGCGGGAGGCGCTCTCCGAGAGCCCGCAAGCTCAG 610
DB 635 GCTGTGAGAGTGGGGCGCTCGGAGCGGGAGGCGCTCTCCGAGAGCCCGCAAGCTCAG 694
QY 611 CCTCATGCGCTGTCTTCCAGGGAATCTCGGGCCAGAGAGCTGTCCAGCAGCTGACGC 670
DB 695 CCTCATGCGCTGTCTTCCAGGGAATCTCGGGCCAGAGAGCTGTCCAGCAGCTGACGC 754
QY 671 TCCCGAGATGGGCTTCTCTGTTGCGGGAGTCCGGGCGCCAGCCCGGGGAGTACGTCT 730
DB 755 TCCCGAGATGGGCTTCTCTGTTGCGGGAGTCCGGGCGCCAGCCCGGGGAGTACGTCT 814
QY 731 GTGCGTGAAGCTTTGGCGCGAGCTCATCCAGCTGTCGACAGCGGAGCGCCACT 790
DB 815 GTGCGTGAAGCTTTGGCGCGAGCTCATCCACTACCGCGTGTGCAACCGGAGCGCCACT 874
QY 791 CACATTCATGAGCGCGCTTCTCTGCAACCTCATGAGCATGTGTGAGCATTTACGCAA 850
DB 875 CACATTCATGAGCGCGCTTCTCTGCAACCTCATGAGCATGTGTGAGCATTTACGCAA 934
QY 851 GGACAAAGGCGCTATGTGCAACAGCTGTGAGAACCAAGCGGAACAGCGGACCAAGTC 910
DB 935 GGACAAAGGCGCTATGTGCAACAGCTGTGAGAACCAAGCGGAACAGCGGACCAAGTC 994
QY 911 GGCCGAGAGAGAGCTGGCCAGGCGGGCTGTACTGAACTGTCAGCATTTGCAATTTGGG 970
DB 995 GGCCGAGAGAGAGCTGGCCAGGCGGGCTGTACTGAACTGTCAGCATTTGCAATTTGGG 1054
QY 971 AACACAGATCGGAGAGGAGAGTGTGAGACTGTCTTCCGAGGGGTGAGTACTTGGGCAAAA 1030
DB 1055 AACACAGATCGGAGAGGAGAGTGTGAGACTGTCTTCCGAGGGGTGAGTACTTGGGCAAAA 1114
QY 1031 GGTGGCGCTGAGAAATATCAAGTGTGATGTGACAGGCGCGGCTTCTGAGAGAGCGGC 1090
DB 1115 GGTGGCGCTGAGAAATATCAAGTGTGATGTGACAGGCGCGGCTTCTGAGAGAGCGGC 1174
QY 1091 CGTCATGAGAGAGATGCAACAGAGAACTGTGCGCTCTCTGAGGCGGTATCCTGACCA 1150
DB 1175 CGTCATGAGAGAGATGCAACAGAGAACTGTGCGCTCTCTGAGGCGGTATCCTGACCA 1234
QY 1151 GGGGCTGTACATTTGTATGAGACAGTGTGAGCAAGGGCAACCTGTGTCTCTGCGGAC 1210
DB 1235 GGGGCTGTACATTTGTATGAGACAGTGTGAGCAAGGGCAACCTGTGTCTCTGCGGAC 1294

```

```

QY 1211 CCGGGGTGAGCCCTGTGTAACACCGCTCAGCTCAGATTTTCTCTGACGTGGCCA 1270
DB 1295 CCGGGGTGAGCCCTGTGTAACACCGCTCAGCTCAGATTTTCTCTGACGTGGCCA 1354
QY 1271 GGGCATGAGTACCTTGGAGAGCAAGAGCTTGTGACCGCGAGCCTGGCCCGCAACAT 1330
DB 1355 GGGCATGAGTACCTTGGAGAGCAAGAGCTTGTGACCGCGAGCCTGGCCCGCAACAT 1414
QY 1331 CCTGTCTCAGAGAGACCTGTGGCCAGGTCAAGCTTTGGCTTGGCCCAAGCGAGG 1390
DB 1415 CCTGTCTCAGAGAGACCTGTGGCCAGGTCAAGCTTTGGCTTGGCCCAAGCGAGG 1474
QY 1391 GAAGGGGTGAGACTCAAGCCGCGCTGCGCTCAAGTGTGAGCGGCGCCGAGGCTCAACA 1450
DB 1475 GAAGGGGTGAGACTCAAGCCGCGCTGCGCTCAAGTGTGAGCGGCGCCGAGGCTCAACA 1534
QY 1451 CGGGAAGTTCACAGCAAGTCCGATGTGAGAGTTTGGGGTCTCTCTGGAGAGCTCT 1510
DB 1535 CGGGAAGTTCACAGCAAGTCCGATGTGAGAGTTTGGGGTCTCTCTGGAGAGCTCT 1594
QY 1511 CTCATATGAGAGGCTTCCCTTACCCTTAAATGTCTACTGAAAGAGGTGTGAGAGCTGGA 1570
DB 1595 CTCATATGAGAGGCTTCCCTTACCCTTAAATGTCTACTGAAAGAGGTGTGAGAGCTGGA 1654
QY 1571 GAAAGGGTACCGGATGAGAACCCCGGAGGGCTGTCCAGGCGCCCGTGCAGCTCTCATGAG 1630
DB 1655 GAAAGGGTACCGGATGAGAACCCCGGAGGGCTGTCCAGGCGCCCGTGCAGCTCTCATGAG 1714
QY 1631 CAGCTGCTGGAGAGGAGAGCCCGCCCGCGGCGCACCTTCCGCAAACTGGCGAGAGAGT 1690
DB 1715 CAGCTGCTGGAGAGGAGAGCCCGCCCGCGGCGCACCTTCCGCAAACTGGCGAGAGAGT 1774
QY 1691 GGGCGGGAGCTACGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1750
DB 1775 GGGCGGGAGCTACGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1834
QY 1751 CTCACACTGCGCCCGGAGAGCAGAGAGCTGTGAGAGCCCGCGGAGGGGCTTGGAGCCGAGA 1810
DB 1835 CTCACACTGCGCCCGGAGAGCAGAGAGCTGTGAGAGCCCGCGGAGGGGCTTGGAGCCGAGA 1894
QY 1811 GGAACGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1870
DB 1895 GGAACGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1954
QY 1871 CGGGCAAGTATCTCTCTGTGCTGCGCCAGCAGAGGGGTGTGAGAGTGTGAGAGTGTGAGAG 1930
DB 1955 CGGGCAAGTATCTCTCTGTGCTGCGCCAGCAGAGGGGTGTGAGAGTGTGAGAGTGTGAGAG 2014
QY 1931 GGCCCGTGGAGACCCCGAGAGCTGTGAGAGATGTGCGCGATTAAGAGAGATTTCTAAGG 1989
DB 2015 GGCCCGTGGAGACCCCGAGAGCTGTGAGAGATGTGCGCGATTAAGAGAGATTTCTAAGG 2073

```

```

RESULT 14
S71669 1744 bp mRNA linear PRI 24-OCT-1994
LOCUS S71669
DEFINITION Isk-leukocyte carboxyl-terminal src kinase related gene [human],
peripheral T cells, mRNA, 1744 nt.
ACCESSION S71669
VERSION S71669.1 GI:559593
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1744)
AUTHORS McVicar,D.W., Lai,B.K., Lloyd,A., Kawamura,M., Chen,Y.Q., Zhang,X.,
Staples,J.E., Ortaldo,J.R. and O'Shea,J.J.
TITLE Molecular cloning of Isk, a carboxyl-terminal src kinase (csk)
related gene, expressed in leukocytes
JOURNAL Oncogene 9 (7), 2037-2044 (1994)
MEDLINE 94268844

```

PUBMED 7516063
GenBank staff at the National Library of Medicine created this entry [NCBI g15068] from the original journal article.
This sequence comes from Fig. 1.

FEATURES
source
1.1744
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1.1744
/gene="lsk"
/note="leukocyte carboxyl-terminal src kinase related gene"

gene

CDS

/note="leukocyte carboxyl-terminal src kinase related gene"
125..1522
/gene="lsk"
/note="leukocyte carboxyl-terminal src kinase related gene: This sequence comes from Fig. 1"
/codon_start=1
/protein_id="AAB30995.1"
/db_xref="GI:559594"
/translation="MPTRMAPGTQCTICENTRPKPELAFKGDVVTILEACENKS
WYVKHHTSGOEGILLAGALREERLADPKLSIMPFHGISGOAVOOLPEDEGL
FLVRESGHPDVLCTVGRDVIHYRVLRHDSHTIDEVFPNCIMDAVEHRSKRG
AICOTKLVRKRRKHTKSABEELARAGMLNLOHTLGAQIGSEBNAVIOGHTLGRK
VNRKCDVTAQAFLEDTAVTKMHNLYRLGVIIIDGILYMEHVSNGENLFLRT
RGALVNTAQLQFSLHREGMETLESKLVHRDLAARNLIVSEELAVVSDPLAKA
ERKGLDSRLPYKMTAPBALKHGKFTSKSDVMSFGLWEVEFSGRAPPKMLKEVS
EAEKRYRMEPEPGCPYHVLMSCEMEAPARRPFRKLAERLSAGAPASVS
GQADGSTPSRSEP"

BASE COUNT 354 a 534 c 582 g 274 t
ORIGIN

Query Match 82.6%; Score 1651.4; DB 9; Length 1744;
Best Local Similarity 98.3%; Pred. No. 3.8e-237;
Matches 1712; Conservative 0; Mismatches 21; Indels 8; Gaps 4;

QY 260 GGGGGGGGAGGCTCTCTGTTCTCTGGGGGGGAGGATTCACGCTTGATTCCTAGGA 319
DB 4 GGGGGGGGAGGCTCTCTGTTCTCTGGGGGGGAGGATTCACGCTTGATTCCTAGGA 63
QY 320 ACTTCCCGGGTGGAGCCCGCTTCTCCGAGCCTGGACACCCCTCCGCTCAGCAG 379
DB 64 ACTTCCCGGGTGGAGCCCGCTTCTCCGAGCCTGGACACCCCTCCGCTCAGCAG 123
QY 380 GATGCCAAGAGGCGTGGGCCCCGGGCAACCAAGTATCACCAATGGAGCACACCG 439
DB 124 GATGCCAAGAGGCGTGGGCCCCGGGCAACCAAGTATCACCAATGGAGCACACCG 183
QY 440 CCCCAGCCAGGAGGAGCTGGCTTCCGCAAGGCGACGTGTACCATCTCTGAGGCTG 499
DB 184 CCCCAGCCAGGAGGAGCTGGCTTCCGCAAGGCGACGTGTACCATCTCTGAGGCTG 243
QY 500 CGAAGCAAGAGCTGTACCGCGTCAAGCAACACAGAGTGGAGAGGAGGAGTGGG 559
DB 244 CGAAGCAAGAGCTGTACCGCGTCAAGCAACACAGAGTGGAGAGGAGGAGTGGG 303
QY 560 AGCTGGGCGCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 619
DB 304 AGCTGGGCGCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 363
QY 620 GTGGTTCACAGGAAATCTCGGCGCAGAGAGGCTGTCCAGACAGTCAAGCTCCAGGA 679
DB 364 GTGGTTCACAGGAAATCTCGGCGCAGAGAGGCTGTCCAGACAGTCAAGCTCCAGGA 423
QY 680 TGGGCTGTCTCTGTCGCGAGAGTCCGCGGCGCACCCCGGAGTACGTCCTGTCGTAG 739
DB 424 TGGGCTGTCTCTGTCGCGAGAGTCCGCGGCGCACCCCGGAGTACGTCCTGTCGTAG 483
QY 740 CTTTGGCGCGAGCTATCACCTACCGCTGTGTGACCGCGAGAGGCGCACCTCAATCGA 799
DB 484 GTTTGGCGCGAGCTATCACCTACCGCTGTGTGACCGCGAGAGGCGCACCTCAATCGA 543
QY 800 TGAAGCCGCTGTTCTTTCGACACCTATGACATGTGTGAGCATTTACAGCAGACAAAGG 859

DB 544 TGAAGCCGCTGTTCTTCTGCAACCTCATGTGACATGTGTGAGCATTTACAGCAGACAAAGG 603
QY 860 CGTATCTGACACCACTGGTGTGAGCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 919
DB 604 CGCTATCTGACCACTGGTGTGAGCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 663
QY 920 GGAAGTGGCGAGGCGGGGCTGTACTGAACTGTGACAGATTGTGACATTGGAGACAGAT 979
DB 664 GGAAGTGGCGAGGCGGGGCTGTACTGAACTGTGACAGATTGTGACATTGGAGACAGAT 723
QY 980 CGAGAGGAGAGATTGGAGCTTCTGCAAGGCTGATACCTGTGGGCAAAAGTGCCGT 1039
DB 724 CGAGAGGAGAGATTGGAGCTTCTGCAAGGCTGATACCTGTGGGCAAAAGTGCCGT 780
QY 1040 GAAGATATCAATGTGTGATGTGACAGCCAGGCTTCTGAGAGAGAGGAGGAGGAGGAGG 1099
DB 781 GAAGATATCAATGTGTGATGTGACAGCCAGGCTTCTGAGAGAGAGGAGGAGGAGGAGG 840
QY 1100 GAAGATGCAACAGAGAACCTGTGTGCTCTGAGGCGTGTATCTGACAGGAGGCTGTA 1159
DB 841 GAAGATGCAACAGAGAACCTGTGTGCTCTGAGGCGTGTATCTGACAGGAGGCTGTA 900
QY 1160 CATTGTGATGAGACAGTGTGAGCAAGGCAACCTGTGTGACTTCTGCGGAGCCGGGGTGC 1219
DB 901 CATTGTGATGAGACAGTGTGAGCAAGGCAACCTGTGTGACTTCTGCGGAGCCGGGGTGC 960
QY 1220 AGCCCTGTGAAACAGCGCTCAGCTCTGCAAGTTTCTGACAGCTGGCGAGGAGGAGGAG 1279
DB 961 AGCCCTGTGAAACAGCGCTCAGCTCTGCAAGTTTCTGACAGCTGGCGAGGAGGAGGAG 1020
QY 1280 GTACTGTGAGAGCAAGAGCTGTGTGACAGCGAGCTGCGCGCCGCAACATCTGTCTC 1339
DB 1021 GTACTGTGAGAGCAAGAGCTGTGTGACAGCGAGCTGCGCGCCGCAACATCTGTCTC 1080
QY 1340 AGAGGACCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1399
DB 1081 AGAGGAGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
QY 1400 AGACTCAAGCGGCTGCTGCTCAAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1459
DB 1141 AGACTCAAGCGGCTGCTGCTCAAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
QY 1460 CACGACAGTGGATGTGTGAGTTTGGGAGTGTCTGTGTGAGGAGTCTTTCATATGG 1519
DB 1201 CACGACAGTGGATGTGTGAGTTTGGGAGTGTCTGTGTGAGGAGTCTTTCATATGG 1260
QY 1520 ACGGCTCGCTAACCTTAAATGTCACTGAAAGAGTGTGTGAGGAGGAGGAGGAGGAGG 1579
DB 1261 ACGGCTCGCTAACCTTAAATGTCACTGAAAGAGTGTGTGAGGAGGAGGAGGAGGAGG 1320
QY 1580 CCGCATGGAACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1639
DB 1321 CCGCATGGAACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
QY 1640 GGAAGCAGAGCCCGCGCGGCGCACCTTTCGCAAACTGTGCGGAGAGAGCTGGCCGGA 1699
DB 1381 GGAAGCAGAGCCCGCGCGGCGCACCTTTCGCAAACTGTGCGGAGAGAGCTGGCCGGA 1440
QY 1700 GCTAGCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1759
DB 1441 GCTAGCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
QY 1760 GCCCGCAAGCAGAGAGCTGTGACCCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1819
DB 1501 GCCCGCAAGCAGAGAGCTGTGACCCAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1558
QY 1820 GAGTGAAGAGTGGGCGGTGTGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1879
DB 1559 GAGTGAAGAGTGGGCGGTGTGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1615
QY 1880 CATCTCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1939

Db 1616 CATCTCTCTGTGCTCCACAGCAGGAGGCTGCGCCACGATGAGGGGCTCTGGCGGCCCGCTGG 1675
QY 1940 ACACCCAGACCTCGAGAGATGATCGCCCATTAAGACGATCTTAAGACCTTAATAA 1999
Db 1676 ACACCCAGACCTCGAGAGATGATCGCCCATTAAGAGATCTTAAGAGAAAACAAA 1735
QY 2000 A 2000
Db 1736 A 1736

RESULT 15
AR053291 1521 bp DNA Linear PAT 29-SEP-1999
LOCUS AR053291
DEFINITION Sequence 10 from patent US 5834208.
ACCESSION AR053291
VERSION AR053291.1 GI:5978153
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1521)
AUTHORS Sakano, S.
TITLE Tyrosine kinase
JOURNAL Patent: US 5834208-A 10 10-NOV-1998;
FEATURES
source 1. 1521
/organism="unknown"
BASE COUNT 295 a 470 c 507 g 249 t
ORIGIN

Query Match 76.0%; Score 1519.4; DB 6; Length 1521;
Best Local Similarity 99.9%; Pred. No. 2e-217;
Matches 1520; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 258 ATGGGGGGGCGAGGCTCTGTGTTCTGTGGGGGCAATTTCAGGGCTGATTCGCTAG 317
Db 1 ATGGGGGGGCGAGGCTCTGTGTTCTGTGGGGGCAATTTCAGGGCTGATTCGCTAG 60
QY 318 GAACCTCCCGGGGTAGAGCCCGCTCTCGAGGCTGGACCCCGCTCCGCTCAGCC 377
Db 61 GAACCTCCCGGGGTAGAGCCCGCTCTCGAGGCTGGACCCCGCTCCGCTCAGCC 120
QY 378 AGAATGCCAAGAGGCGCTGGGCGCCGGGCAACCGATATACCAATGAGGACACC 437
Db 121 AGAATGCCAAGAGGCGCTGGGCGCCGGGCAACCGATATACCAATGAGGACACC 180
QY 438 CGCCCAAGCAGAGGAGCTGGCTTCCGCAAGGGGCACTGTGCATCCTGAGGCC 497
Db 181 CGCCCAAGCAGAGGAGCTGGCTTCCGCAAGGGGCACTGTGCATCCTGAGGCC 240
QY 498 TGGCGAACAAGAGTGTACCGCTCAAGCAACACAGAGGAGAGGGGCTGTG 557
Db 241 TGGCGAACAAGAGTGTACCGCTCAAGCAACACAGAGGAGAGGGGCTGTG 300
QY 558 GCACCTGGGGGCTCGGGAGCGAGGAGCCCTCTCCGAGACCCCAAGCTCAGCTCATG 617
Db 301 GCACCTGGGGGCTCGGGAGCGAGGAGCCCTCTCCGAGACCCCAAGCTCAGCTCATG 360
QY 618 CCGTGTTCACAGGAGATCTGGGCGACAGAGGCTGTCCAGACCTCAGCTCCGAG 677
Db 361 CCGTGTTCACAGGAGATCTGGGCGACAGAGGCTGTCCAGACCTCAGCTCCGAG 420
QY 678 GATGGGCTGCTCCTGTGGGAGTCCGCGCCGAGGAGGAGGAGTCTGAGGCTG 737
Db 421 GATGGGCTGCTCCTGTGGGAGTCCGCGCCGAGGAGGAGGAGTCTGAGGCTG 480
QY 738 AGCTTTGGCCGCGAGCTATCCATACGAGCTGTGACACCGGAGGCGACCTCAATC 797
Db 481 AGCTTTGGCCGCGAGCTATCCATACGAGCTGTGACACCGGAGGCGACCTCAATC 540
QY 798 GATGAGGCGGCTTTCTTCTGCACTCATGACATGATGAGCATTTACAGCAAGACAAG 857

Db 541 GATGAGGCGGCTTTCTTCTGCACTCATGACATGATGAGCATTTACAGCAAGACAAG 600
QY 858 GGCCTATCTGCTACCAACCTGTGTAGACCAAGGAGAAACAGGGACCAATCGCCGAG 917
Db 601 GGCCTATCTGCTACCAACCTGTGTAGACCAAGGAGAAACAGGGACCAATCGCCGAG 660
QY 918 GAGGAGCTGGCCAGGCGGCTGTACTGAACTGACAGATTGACATTGGAGACAG 977
Db 661 GAGGAGCTGGCCAGGCGGCTGTACTGAACTGACAGATTGACATTGGAGACAG 720
QY 978 ATCGAGAGGAGAGATTGAGCTGTCTGCAAGGCTGATACCTGAGGCAAAAGTGCC 1037
Db 721 ATCGAGAGGAGAGATTGAGCTGTCTGCAAGGCTGATACCTGAGGCAAAAGTGCC 780
QY 1038 GTGAAGATATCAAGTGTGATGATACAGCCAGGCTCTCGAGAGAGAGGCGTCAAG 1097
Db 781 GTGAAGATATCAAGTGTGATGATACAGCCAGGCTCTCGAGAGAGAGGCGTCAAG 840
QY 1098 ACGAAGATGCAACAGAGACCTGTGGCTCTCGTGGGCTGATCTGACACAGGAGCTG 1157
Db 841 ACGAAGATGCAACAGAGACCTGTGGCTCTCGTGGGCTGATCTGACACAGGAGCTG 900
QY 1158 TACATTTGATGAGACAGTGTAGCAAGAGGCAACCTGTGTAACCTTTCGGAACCCGGGT 1217
Db 901 TACATTTGATGAGACAGTGTAGCAAGAGGCAACCTGTGTAACCTTTCGGAACCCGGGT 960
QY 1218 CGAGCCCTGTAACACCGCTCAGCTCCTGTGAGTTTCTGACAGTGGCGAGGAGCTG 1277
Db 961 CGAGCCCTGTAACACCGCTCAGCTCCTGTGAGTTTCTGACAGTGGCGAGGAGCTG 1020
QY 1278 GAGTACTGTGAGAGCAAGACCTGTGACAGGCGCACTGAGGCGCCGCAACATCTGTGTC 1337
Db 1021 GAGTACTGTGAGAGCAAGACCTGTGACAGGCGCACTGAGGCGCCGCAACATCTGTGTC 1080
QY 1338 TCAGAGGACCTGTGTGGCCAGAGTGTGAGCACTTGTGGCCAAAGCCGAGAGGGG 1397
Db 1081 TCAGAGGACCTGTGTGGCCAGAGTGTGAGCACTTGTGGCCAAAGCCGAGAGGGG 1140
QY 1398 CTAGACTCAAGCCGGCTGCGCTCAAGAGGAGGCGCCGAGGCTCTCAACACAGGGAG 1457
Db 1141 CTAGACTCAAGCCGGCTGCGCTCAAGAGGAGGCGCCGAGGCTCTCAACACAGGGAG 1200
QY 1458 TTCACACAGCAAGTGTGATGTGTGAGTTTGGGGTGTGCTCTGTGAGGCTCTTCATAT 1517
Db 1201 TTCACACAGCAAGTGTGATGTGTGAGTTTGGGGTGTGCTCTGTGAGGCTCTTCATAT 1260
QY 1518 GGAGCGGCTCCGTACCTTAATAATGTCACTGAAGAAGTGTGGAGGCCGTGAGAAAGGG 1577
Db 1261 GGAGCGGCTCCGTACCTTAATAATGTCACTGAAGAAGTGTGGAGGCCGTGAGAAAGGG 1320
QY 1578 TACCGCATGGAACCCCGGAGGCTGTCCAGGCGCCGTCAGCTCTCATGAGCAGCTGC 1637
Db 1321 TACCGCATGGAACCCCGGAGGCTGTCCAGGCGCCGTCAGCTCTCATGAGCAGCTGC 1380
QY 1638 TGGGAGGAGAGCCCGCCGCGGCGCACCTTCCGCAAACTGAGCCGAGAAAGCTGGCCGG 1697
Db 1381 TGGGAGGAGAGCCCGCCGCGGCGCACCTTCCGCAAACTGAGCCGAGAAAGCTGGCCGG 1440
QY 1698 GAGCTACGATGCAAGGTGTCACAGCTTCCTGTAGAGGAGAGAGCCGAGGCTGCAC 1757
Db 1441 GAGCTACGATGCAAGGTGTCACAGCTTCCTGTAGAGGAGAGAGCCGAGGCTGCAC 1500
QY 1758 TCGCCCGGAGGAGGAGGCC 1778
Db 1501 TCGCCCGGAGGAGGAGGCC 1521

Search completed: August 1, 2003, 18:16:28
Job time : 9498 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 15:19:49 ; Search time 182 Seconds
(Without alignments) 4850.364 Million cell updates/sec

Title: US-09-977-260-1
Perfect score: 2000
Sequence: 1 ctgcctccaagtgtgtgcagc.....attctaagactctaaanaa 2000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/jna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/jna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/jna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/jna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/jna/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/jna/Backfillseq.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Query	Length	DB	ID	Description
1	2000	100.0	2000	4	US-08-426-509A-1			Sequence 1, Appli
2	2000	100.0	2000	4	US-08-232-545-1			Sequence 1, Appli
3	2000	100.0	2000	5	PCT-US95-05008-1			Sequence 1, Appli
4	1909.8	95.5	1987	2	US-08-876-882-1			Sequence 1, Appli
5	1909.8	95.5	1987	4	US-09-315-928-1			Sequence 1, Appli
6	1903.4	95.2	1942	2	US-08-604-989A-11			Sequence 1, Appli
7	1519.4	76.0	1521	2	US-08-604-989A-10			Sequence 10, Appli
8	1398	69.9	1398	2	US-08-604-989A-9			Sequence 1, Appli
9	1377	68.8	1713	4	US-09-741-154-1			Sequence 9, Appli
10	738	36.9	738	4	US-08-604-989A-8			Sequence 1, Appli
11	455	22.8	16389	4	US-09-741-154-3			Sequence 3, Appli
12	225	11.2	225	1	US-08-604-989A-7			Sequence 7, Appli
13	212.4	10.6	1611	1	US-07-820-011A-3			Sequence 3, Appli
14	212.4	10.6	1611	5	PCT-US93-00445-3			Sequence 3, Appli
15	199.2	10.0	1602	1	US-07-820-011A-1			Sequence 1, Appli
16	199.2	10.0	1602	5	PCT-US93-00445-1			Sequence 1, Appli
17	192	9.6	192	2	US-08-604-989A-6			Sequence 6, Appli
18	185	9.2	1574	3	US-09-173-581-12			Sequence 12, Appli
19	185	9.2	1574	3	US-09-420-915-12			Sequence 12, Appli
20	181.8	9.1	3623	1	US-08-306-691B-35			Sequence 35, Appli
21	170.4	8.5	1467	4	US-09-579-182-2			Sequence 2, Appli
22	168.8	8.4	1548	4	US-09-099-053-1			Sequence 1, Appli
23	165.8	8.3	2674	4	US-09-817-180-1			Sequence 1, Appli
24	162	8.1	5993	3	US-09-383-630-1			Sequence 1, Appli
25	162	8.1	5993	3	US-09-383-630-2			Sequence 2, Appli
26	151.2	7.6	2647	4	US-09-220-132-77			Sequence 77, Appli
27	151.2	7.6	2647	5	PCT-US93-06251-77			Sequence 77, Appli

28	148.8	7.4	2049	4	US-09-099-749-10			Sequence 10, Appli
29	148.8	7.4	2493	4	US-09-620-312D-830			Sequence 830, App
30	148.8	7.4	2598	4	US-09-417-197-110			Sequence 110, App
31	148.8	7.4	2616	4	US-09-417-197-108			Sequence 108, App
32	148	7.4	2469	1	US-08-459-286-1			Sequence 1, Appli
33	146.4	7.3	933	2	US-08-701-191A-4			Sequence 4, Appli
34	146.4	7.3	1056	2	US-08-701-191A-5			Sequence 5, Appli
35	146.4	7.3	2469	1	US-07-997-133-2			Sequence 2, Appli
36	146.4	7.3	2469	5	US-07-997-133-2			Sequence 2, Appli
37	146.4	7.3	2662	2	US-08-451-822A-14			Sequence 14, Appli
38	146.4	7.3	2662	2	US-08-323-430-14			Sequence 14, Appli
39	146.4	7.3	2733	1	US-08-371-001-14			Sequence 14, Appli
40	146.4	7.3	2733	5	PCT-US96-00331-14			Sequence 14, Appli
41	137.6	6.9	1491	2	US-09-006-675-1			Sequence 1, Appli
42	137.6	6.9	1491	3	US-09-228-603A-1			Sequence 1, Appli
43	136.6	6.8	3945	4	US-09-016-434-1404			Sequence 1404, Ap
44	136.6	6.8	3969	1	US-08-436-044-5			Sequence 5, Appli
45	136.6	6.8	3969	1	US-08-222-616-23			Sequence 23, Appli

ALIGNMENTS

RESULT 1
US-08-426-509A-1
Sequence 1, Application US/08426509A
Patent No. 6326469
GENERAL INFORMATION:
APPLICANT: Ullrich, Axel
APPLICANT: Gishizky, Mikhail
APPLICANT: Sures, Irman G.
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
TITLE OF INVENTION: TYROSINE KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York,
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,509A
FILING DATE: 21-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/232,545
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 7683-0074-9999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
US-08-426-509A-1
Query Match 100.0% Score 2000; DB 4; Length 2000;
Best Local Similarity 100.0% Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGCTCCAAAGTTGTGACGCGGAGCCGCTCGGGGTGTGACAGCGGGCTCGGAGAGCC 60
 Db 1 CTCGCTCCAAAGTTGTGACGCGGAGCCGCTCGGGGTGTGACAGCGGGCTCGGAGAGCC 60
 QY 61 TCCTGG 120
 Db 61 TCCTGG 120
 QY 121 AGGCTGAGTCCAGTGGGACCCAGCTCCCTACCTGCTGTGACAGCGGGCTGGCTGTGCA 180
 Db 121 AGGCTGAGTCCAGTGGGACCCAGCTCCCTACCTGCTGTGACAGCGGGCTGGCTGTGCA 180
 QY 181 GGGCATTCCAGGGTCCCGGACAGTGTACCACTGTGCTGAGTGTCTCTGCTGCTGCTGAG 240
 Db 181 GGGCATTCCAGGGTCCCGGACAGTGTACCACTGTGCTGAGTGTCTCTGCTGCTGCTGAG 240
 QY 241 TTTCCTCTGTGG 300
 Db 241 TTTCCTCTGTGG 300
 QY 301 GCTGTGATTTCTGTGAGAACTTCCCGGGTGAAGCCCGCTTCTCCAGGCTGTGACCC 360
 Db 301 GCTGTGATTTCTGTGAGAACTTCCCGGGTGAAGCCCGCTTCTCCAGGCTGTGACCC 360
 QY 361 CCCCTCCCGTCTGAGGAGATGCAACGAGGGGCTGTGGCCCGGGGACCCAGTGTATCA 420
 Db 361 CCCCTCCCGTCTGAGGAGATGCAACGAGGGGCTGTGGCCCGGGGACCCAGTGTATCA 420
 QY 421 CCAAATGCGAGCACACCCCGCCCAAGCGAGGGAGGCTTCCGGAAGGGGAGAGTGG 480
 Db 421 CCAAATGCGAGCACACCCCGCCCAAGCGAGGGAGGCTTCCGGAAGGGGAGAGTGG 480
 QY 481 TCACCATCTCTGAGAGGCTGTGAGAACAGAGTGTATCCGCTCAAGCAACACACAGT 540
 Db 481 TCACCATCTCTGAGAGGCTGTGAGAACAGAGTGTATCCGCTCAAGCAACACACAGT 540
 QY 541 GAGAGAGAGGGGCTGTGAGAGCTGTGGGGGCTGTGGGGAGGGGAGGGCTTCCGAGACC 600
 Db 541 GAGAGAGAGGGGCTGTGAGAGCTGTGGGGGCTGTGGGGAGGGGAGGGCTTCCGAGACC 600
 QY 601 CCAGAGCTCAAGCTCATGAGCTGTGCTCAGCGGAGAACTTCGGGCGAGGAGGCTGTGACG 660
 Db 601 CCAGAGCTCAAGCTCATGAGCTGTGCTCAGCGGAGAACTTCGGGCGAGGAGGCTGTGACG 660
 QY 661 AGCTGAGGCTTCCGAGAGATGGGCTTCTGTGGGGAGTCCGGCGCCACCCCGGCG 720
 Db 661 AGCTGAGGCTTCCGAGAGATGGGCTTCTGTGGGGAGTCCGGCGCCACCCCGGCG 720
 QY 721 ACTAGCTCTGTGGGAGCTTGTGGCGGAGCTCATCCACTACCGGCTGTGACCGG 780
 Db 721 ACTAGCTCTGTGGGAGCTTGTGGCGGAGCTCATCCACTACCGGCTGTGACCGG 780
 QY 781 ACAGGCACTCAACATCATGAGGCGGTGTCTTCTGCAACCTCATGAGCATGTGTGAGC 840
 Db 781 ACAGGCACTCAACATCATGAGGCGGTGTCTTCTGCAACCTCATGAGCATGTGTGAGC 840
 QY 841 ATTACAGCAAGCAAGGCGGCTATCTGACCAAGCTGTGAGAACCAAAAGCGAAACAG 900
 Db 841 ATTACAGCAAGCAAGGCGGCTATCTGACCAAGCTGTGAGAACCAAAAGCGAAACAG 900
 QY 901 GGAACCAAGTGGCGGAGAGAGAGCTGTGCAAGGCGGGCTGTACTGAACTGTGACAGATT 960
 Db 901 GGAACCAAGTGGCGGAGAGAGAGCTGTGCAAGGCGGGCTGTACTGAACTGTGACAGATT 960
 QY 961 TGACATTGGGAGACAGATCGGAGAGGAGAGTGTGAGAGCTGTGCAAGGCTGTGAGTACC 1020
 Db 961 TGACATTGGGAGACAGATCGGAGAGGAGAGTGTGAGAGCTGTGCAAGGCTGTGAGTACC 1020
 QY 1021 TGGGGCAAAAGGTGGCGCTGAAGATATCAAGTGTGATGTGACAGCCGAGGCTTCTCTG 1080
 Db 1021 TGGGGCAAAAGGTGGCGCTGAAGATATCAAGTGTGATGTGACAGCCGAGGCTTCTCTG 1080
 QY 1081 AGGAGACGGCGCTCATGACGAAGATGCAACAGAGAACTGGTGTCTCTCTGGGCGTGA 1140

Db 1081 ACAGAGACGGCGCTCATGACGAGAGATGCAACAGAGAACTGTGTCTCTGTGGGCTGA 1140
 QY 1141 TCCTGACACAGGGGGCTGTACATGTGATGAGAGACGTGAGCAAGGGCAACCTGTGTAAT 1200
 Db 1141 TCCTGACACAGGGGGCTGTATATTGTATGAGAGACGTGAGCAAGGGCAACCTGTGTAAT 1200
 QY 1201 TTCTGGGAGACCCGGGGTTCAGAGCCCTGTGTGAACACCCCTGAGCTCTGCAAGTTTCTCTG 1260
 Db 1201 TTCTGGGAGACCCGGGGTTCAGAGCCCTGTGTGAACACCCCTGAGCTCTGCAAGTTTCTCTG 1260
 QY 1261 ACCTGGCCGAGGAGATGAGTACCTGTGAGAGCAAGAACTTGTGACACCGGACCTGTGCGG 1320
 Db 1261 ACCTGGCCGAGGAGATGAGTACCTGTGAGAGCAAGAACTTGTGACACCGGACCTGTGCGG 1320
 QY 1321 CCCGCAACATCTGTGCTGTGAGAGACCTGTGTGAGCAAGGTCACAGCTTGTGGCTGGCA 1380
 Db 1321 CCCGCAACATCTGTGCTGTGAGAGACCTGTGTGAGCAAGGTCACAGCTTGTGGCTGGCA 1380
 QY 1381 AAGCCGAGGAGGAGGAGCTAGACTCAAGCCGCTGCGCTCAAGTGGAGGCGCCGAGG 1440
 Db 1381 AAGCCGAGGAGGAGGAGCTAGACTCAAGCCGCTGCGCTCAAGTGGAGGCGCCGAGG 1440
 QY 1441 CTCTCAACACGAGGAGTTCACCAAGTGTGAGTGTGTGAGTGTGTGGGGTCTCTCT 1500
 Db 1441 CTCTCAACACGAGGAGTTCACCAAGTGTGAGTGTGTGAGTGTGTGGGGTCTCTCT 1500
 QY 1501 GGGAGGCTTCTATATGAGAGGCTTCCGAGCTTCCGAAATGTATCTGAGAAAGAGTGTGCG 1560
 Db 1501 GGGAGGCTTCTATATGAGAGGCTTCCGAGCTTCCGAAATGTATCTGAGAAAGAGTGTGCG 1560
 QY 1561 AAGCCGTGTGAGAGGAGGATACCGATGAAACCCCGGAGGAGTGTGAGGCGCCGCTGACAG 1620
 Db 1561 AAGCCGTGTGAGAGGAGGATACCGATGAAACCCCGGAGGAGTGTGAGGCGCCGCTGACAG 1620
 QY 1621 TCCTCATGAGCACTGTGTGAGAGGAGAGGCGCCGCGGCGCACCTTCCGAAACTGG 1680
 Db 1621 TCCTCATGAGCACTGTGTGAGAGGAGAGGCGCCGCGGCGCACCTTCCGAAACTGG 1680
 QY 1681 CCGAGAGGCTGGGCGCGGAGCTACGAGTGTGAGAGTGTGAGGCTTCCGAGCTGTGAGGAGG 1740
 Db 1681 CCGAGAGGCTGGGCGCGGAGCTACGAGTGTGAGAGTGTGAGGCTTCCGAGCTGTGAGGAGG 1740
 QY 1741 ACAGGCAAGGCTTCCAGCTTCCGCGGAGAGGAGGCTGTGAGCCACCCGAGGAGGCGCT 1800
 Db 1741 ACAGGCAAGGCTTCCAGCTTCCGCGGAGAGGAGGCTGTGAGCCACCCGAGGAGGCGCT 1800
 QY 1801 TGGCCCCAGAGAGACAGAGAGTGTGAGAGTGTGGGGCTGTGAGGAGTGTGAGGAGGAGG 1860
 Db 1801 TGGCCCCAGAGAGACAGAGAGTGTGAGAGTGTGGGGCTGTGAGGAGTGTGAGGAGGAGG 1860
 QY 1861 AAGGTCAGAGGAGGAGTCAATCTCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
 Db 1861 AAGGTCAGAGGAGGAGTCAATCTCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
 QY 1921 GGGCTGTGGCGGCGCTGTGAGACACCCGAGCTGTGAGAGAGTGTGAGGAGTGTGAGGAGG 1980
 Db 1921 GGGCTGTGGCGGCGCTGTGAGACACCCGAGCTGTGAGAGAGTGTGAGGAGTGTGAGGAGG 1980
 QY 1981 ATTCTAAGAGCTCTTAAAAA 2000
 Db 1981 ATTCTAAGAGCTCTTAAAAA 2000

RESULT 2
 US-08-232-545-1
 ; Sequence 1, Application US/08232545
 ; Patent No. 6506578
 ; GENERAL INFORMATION:
 ; APPLICANT: Ullrich, Axel
 ; APPLICANT: Gishinsky, Mikhail
 ; APPLICANT: Sures, Irman G.
 ; TITLE OF INVENTION: No. 6506578e1 Megakaryocytic Protein Tyrosine

TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-232-545-1

Query Match 100.0%; Score 2000; DB 4; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGCTCCAGTTGTGACGCGGGGAGCCGCTCGGGGTGTGACGCGCGCTCGCGAGGCGCC 60
DB 1 CTCGCTCCAGTTGTGACGCGGGGAGCCGCTCGGGGTGTGACGCGCGCTCGCGAGGCGCC 60
QY 61 TCCTGGGGGCGGGGCGGGGCGGCTCGGGGGCGCCCTGAGCAGAAAAAGAGAAAC 120
DB 61 TCCTGGGGGCGGGGCGGGGCGGCTCGGGGGCGCCCTGAGCAGAAAAAGAGAAAC 120
QY 121 AGGCTCGGTCCAGTGTGACCCAGCTCCTTACCTCTGTGACAGCGCGCTGGCTGTGCA 180
DB 121 AGGCTCGGTCCAGTGTGACCCAGCTCCTTACCTCTGTGACAGCGCGCTGGCTGTGCA 180
QY 181 GGCATTCGCAAGCGTCCCGGACTGTGACACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 181 GGCATTCGCAAGCGTCCCGGACTGTGACACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 TTTCCCTCTGGGGGCGATGCGGGGCGAGGCTCTGTGTTCTGTGCGGGCATTTTCAG 300
DB 241 TTTCCCTCTGGGGGCGATGCGGGGCGAGGCTCTGTGTTCTGTGCGGGCATTTTCAG 300
QY 301 GCTGTGATTTGTGTAGGAACCTTCCCGGGGTGAGCCCGCGCTTCTCTCCGAGCTGGCAC 360
DB 301 GCTGTGATTTGTGTAGGAACCTTCCCGGGGTGAGCCCGCGCTTCTCTCCGAGCTGGCAC 360
QY 361 CCCCTCCGCTCGACGACGAGATGCCAAGAGGCGGTGGGGCCCGGGGACCCAGTATATCA 420
DB 361 CCCCTCCGCTCGACGACGAGATGCCAAGAGGCGGTGGGGCCCGGGGACCCAGTATATCA 420
QY 421 CCAATGTGAGACACCCCGCCCAAGCCAGAGGAGACTGGCTTCCGCAAGGGGCGAGTGG 480
DB 421 CCAATGTGAGACACCCCGCCCAAGCCAGAGGAGACTGGCTTCCGCAAGGGGCGAGTGG 480
QY 481 TCACCATCTGTGAGGCTGTGAGAAAGAGCTGTGATCCGCTCAAGCACCACACAGTGG 540
DB 481 TCACCATCTGTGAGGCTGTGAGAAAGAGCTGTGATCCGCTCAAGCACCACACAGTGG 540

DB 481 TCACCATCTGTGAGGCTGTGAGAAAGAGCTGTGATCCGCTCAAGCACCACACAGTGG 540
QY 541 GACAGAGGGGTGTGTGACAGTGGGGGCGCTGCGGGAGCGGAGGCGCTTCTCCGAGACC 600
DB 541 GACAGAGGGGTGTGTGACAGTGGGGGCGCTGCGGGAGCGGAGGCGCTTCTCCGAGACC 600
QY 601 CCAAGCTAGGCTCATGCGCGGTGTTCACGGGAATGATCTCGGGACGAGAGGCTGTCCAGC 660
DB 601 CCAAGCTAGGCTCATGCGCGGTGTTCACGGGAATGATCTCGGGACGAGAGGCTGTCCAGC 660
QY 661 AGCTGACGCTCCCGAGGATGGGCTGTCTGTGTTGCGGAGTCCGCGGCCACCCGGCG 720
DB 661 AGCTGACGCTCCCGAGGATGGGCTGTCTGTGTTGCGGAGTCCGCGGCCACCCGGCG 720
QY 721 ACTAGGCTGTGCTGAGCTTTTGGCCCCGACGTCATCTACCTACCGCGCTGTGACCGCG 780
DB 721 ACTAGGCTGTGCTGAGCTTTTGGCCCCGACGTCATCTACCTACCGCGCTGTGACCGCG 780
QY 781 ACGGCACCTCAATCATGATGAGCGCGTGTCTTCTGCAACCTCATGAGCATGTGTGAGC 840
DB 781 ACGGCACCTCAATCATGATGAGCGCGTGTCTTCTGCAACCTCATGAGCATGTGTGAGC 840
QY 841 ATTACGACAAGAGAGGGCGCTATCTGCACCAAGCTGTGAGACCAAGCGGAACACG 900
DB 841 ATTACGACAAGAGAGGGCGCTATCTGCACCAAGCTGTGAGACCAAGCGGAACACG 900
QY 901 GGACCAATCGCGCGAGAGAGAGCTGCGCAGGGGCGGTGTGTACTGAACTGTGACATTT 960
DB 901 GGACCAATCGCGCGAGAGAGAGCTGCGCAGGGGCGGTGTGTACTGAACTGTGACATTT 960
QY 961 TGACATTTGGAGACAGATCGGAGAGAGAGATTGGAGCTGTCTGTGAGGGTGTGATACC 1020
DB 961 TGACATTTGGAGACAGATCGGAGAGAGAGATTGGAGCTGTCTGTGAGGGTGTGATACC 1020
QY 1021 TGGGGCAAAAGGTGGCGCTGAAGAAATATCAAGTGTGATGTGACACCGACGCTTCTG 1080
DB 1021 TGGGGCAAAAGGTGGCGCTGAAGAAATATCAAGTGTGATGTGACACCGACGCTTCTG 1080
QY 1081 ACGAGACGGCGCTCATGACGAAGATGCAACACGAGAACCTGTGTGCTCTGGGCGTGA 1140
DB 1081 ACGAGACGGCGCTCATGACGAAGATGCAACACGAGAACCTGTGTGCTCTGGGCGTGA 1140
QY 1141 TCCTGACACAGGGGCTGTATATTGTCTATGAGACAGTGTGAGCAAGGGCAACCTGTGACT 1200
DB 1141 TCCTGACACAGGGGCTGTATATTGTCTATGAGACAGTGTGAGCAAGGGCAACCTGTGACT 1200
QY 1201 TTTCTGGGAGCCCGGGGCTGAGCCCTGTGAACACCGCTGACGCTCTGCAATTTCTGTC 1260
DB 1201 TTTCTGGGAGCCCGGGGCTGAGCCCTGTGAACACCGCTGACGCTCTGCAATTTCTGTC 1260
QY 1261 ACGTGGCGGAGGGCATGTGAGTACCTGTGAGACCAAGAACTGTGACCGGCAACCTGGGCG 1320
DB 1261 ACGTGGCGGAGGGCATGTGAGTACCTGTGAGACCAAGAACTGTGACCGGCAACCTGGGCG 1320
QY 1321 CCCGCAATCTCTGTCTCAGAGACCTGTGTGGCCAAAGTGTGAGCACTTTGGCTGTGCA 1380
DB 1321 CCCGCAATCTCTGTCTCAGAGACCTGTGTGGCCAAAGTGTGAGCACTTTGGCTGTGCA 1380
QY 1381 AAGCCGAGCGGAAGGGCTTAACCTCAACCCGGCTGCGCTCAAGTGTGACGCGCCGAGG 1440
DB 1381 AAGCCGAGCGGAAGGGCTTAACCTCAACCCGGCTGCGCTCAAGTGTGACGCGCCGAGG 1440
QY 1441 CTCCTCAACACGGGAAGTTCAACGACAGTGGAGTGTGAGTTTGGGGTGTGCTCT 1500
DB 1441 CTCCTCAACACGGGAAGTTCAACGACAGTGGAGTGTGAGTTTGGGGTGTGCTCT 1500
QY 1501 GGGAGGCTTCTCATATGAGAGGGCTCCGTACCCCTAAATGTCACTGAAAGAGTGTGCG 1560
DB 1501 GGGAGGCTTCTCATATGAGAGGGCTCCGTACCCCTAAATGTCACTGAAAGAGTGTGCG 1560
QY 1561 AGGCGGTGAGAAAGGGGTACCGCATGTGAACCCCGCGAGGCTGTCCAGGCCCGTGCAGC 1620
DB 1561 AGGCGGTGAGAAAGGGGTACCGCATGTGAACCCCGCGAGGCTGTCCAGGCCCGTGCAGC 1620

QY 1621 TCCTCATGAGCAGCTGCTGGAGGAGAGCCCGCCGCGCCACCCCTTCCGCAAACTGG 1680
DB 1621 TCCTCATGAGCAGCTGCTGGAGGAGAGCCCGCCGCGCCACCCCTTCCGCAAACTGG 1680
QY 1681 CCGAGAGAGTGGCCCGGGAGCTACGACAGTGCAGAGTGTCCCGACCTTCCGCTACAGGCAAG 1740
DB 1681 CCGAGAGAGTGGCCCGGGAGCTACGACAGTGCAGAGTGTCCCGACCTTCCGCTACAGGCAAG 1740
QY 1741 AGCGGAGAGCTTCCAGCTGCGCCCGAAGAGCCCTGACCCCGAGGCGGTGGGGCCCT 1800
DB 1741 AGCGGAGAGCTTCCAGCTGCGCCCGAAGAGCCCTGACCCCGAGGCGGTGGGGCCCT 1800
QY 1801 TGGCCCGCAGAGAGCCGAGAGTGGAGAGTGGCGGCTGGGGGCACTGACCAAGGCCAAG 1860
DB 1801 TGGCCCGCAGAGAGCCGAGAGTGGAGAGTGGCGGCTGGGGGCACTGACCAAGGCCAAG 1860
QY 1861 AGGCTTCAGGCGGGGCAAGTCACTCTCTGTCGCCCAAGAGGCGCTGGCCCAAGTGG 1920
DB 1861 AGGCTTCAGGCGGGGCAAGTCACTCTCTGTCGCCCAAGAGGCGCTGGCCCAAGTGG 1920
QY 1921 GGCTCTGGGCGCGCCGTGGACACCCAGACCTGCGAAGAGTATGCGCCCGATAAAGAGCG 1980
DB 1921 GGCTCTGGGCGCGCCGTGGACACCCAGACCTGCGAAGAGTATGCGCCCGATAAAGAGCG 1980
QY 1981 ATTCTAAGAGCTCTTAAAAA 2000
DB 1981 ATTCTAAGAGCTCTTAAAAA 2000

RESULT 3
PCT-US95-05008-1
Sequence 1, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissenschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-APR-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2000 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
PCT-US95-05008-1

Query Match 100.0%; Score 2000; DB 5; Length 2000;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCAAAGTTGTGTGACACCGGAGCCGCTCGGGGTGTGACCGGCTCGGGAGGCC 60
DB 1 CTGCTCCAAAGTTGTGTGACACCGGAGCCGCTCGGGGTGTGACCGGCTCGGGAGGCC 60
QY 61 TCTGTGGGGGCGGGCGGGGGGCTGTGGGGGGCCCCCTGTGACAGAAAACAGAAAGAAC 120
DB 61 TCTGTGGGGGCGGGCGGGGGGCTGTGGGGGGCCCCCTGTGACAGAAAACAGAAAGAAC 120
QY 121 AGGCTGGGTCCAGATGGGACACCAAGCTCCCTACCTGTGTGACAGCCGCTGGCTGTGGCA 180
DB 121 AGGCTGGGTCCAGATGGGACACCAAGCTCCCTACCTGTGTGACAGCCGCTGGCTGTGGCA 180
QY 181 GGCATTTCCAGCGTCCCGGAGCTGTGACCACTTGTCTCACTGTGCTTCACTGCTCAG 240
DB 181 GGCATTTCCAGCGTCCCGGAGCTGTGACCACTTGTCTCACTGTGCTTCACTGCTCAG 240
QY 241 TTTCCCTCTGGGGGGCGATGGCGGGGCGAGGCTCTGTGTTTCTGTGGGGGATTTACAG 300
DB 241 TTTCCCTCTGGGGGGCGATGGCGGGGCGAGGCTCTGTGTTTCTGTGGGGGATTTACAG 300
QY 301 GCTGTGATCTCTGTGAGAACTTCCCGGGGTGAGCCCGCGCTTCTCCGAGGCTGGACAC 360
DB 301 GCTGTGATCTCTGTGAGAACTTCCCGGGGTGAGCCCGCGCTTCTCCGAGGCTGGACAC 360
QY 361 CCCCTCCGCTCTCAGCCAGATGCCAAGAGGCGCTGGGCCCGGGGACCCAGTATCA 420
DB 361 CCCCTCCGCTCTCAGCCAGATGCCAAGAGGCGCTGGGCCCGGGGACCCAGTATCA 420
QY 421 CCNAATGGAGACACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGCGAGTGG 480
DB 421 CCNAATGGAGACACACCCGCCCAAGCCAGGGGAGCTGGCTTCCGCAAGGGCGAGTGG 480
QY 481 TCACCATCTGTGAGGCTCGAGAACAGAGCTGTGACCGGCTCAAGCACACACAGTG 540
DB 481 TCACCATCTGTGAGGCTCGAGAACAGAGCTGTGACCGGCTCAAGCACACACAGTG 540
QY 541 GACAGAGGGGCTGTGAGCTGTGGGCGCTGCGGAGCGGAGGCGCCCTTCGCGAGACC 600
DB 541 GACAGAGGGGCTGTGAGCTGTGGGCGCTGCGGAGCGGAGGCGCCCTTCGCGAGACC 600
QY 601 CCAAGCTCAGCTCATGCCGTGTTCCACNGGGAATATCTCGGGGCGAGAGGCTGTGCCAGC 660
DB 601 CCAAGCTCAGCTCATGCCGTGTTCCACNGGGAATATCTCGGGGCGAGAGGCTGTGCCAGC 660
QY 661 ACCTCAGGCTCCCAAGATGGGCTGTCTCGTGGGAGTCCGCGCGCACCCCGGGG 720
DB 661 ACCTCAGGCTCCCAAGATGGGCTGTCTCGTGGGAGTCCGCGCGCACCCCGGGG 720
QY 721 ACTAGTCTGTGCTGTGAGCTTTGGCCGCGAGCTATCCATACAGCGCTGTGACCGCG 780
DB 721 ACTAGTCTGTGCTGTGAGCTTTGGCCGCGAGCTATCCATACAGCGCGTGTGACCGCG 780
QY 781 AGGCGCAGCTCAGATTCATGAGCGCGGTCTTCTGTGACACTCATGACATGTGTGAGC 840
DB 781 AGGCGCAGCTCAGATTCATGAGCGCGGTCTTCTGTGACACTCATGACATGTGTGAGC 840
QY 841 ATTACAGCAAGAGCAAGGCGCTATCTGACCAAGCTGTGTGAGCAAAAGCGAAACAG 900
DB 841 ATTACAGCAAGAGCAAGGCGCTATCTGACCAAGCTGTGTGAGCAAAAGCGAAACAG 900

QY 901 GGACCAAGTGGCCGAGAGAGAGCTGGCCAGGCGGCTGGTTACTGAACTGCAGCATT 960
 DB 901 GGACCAAGTGGCCGAGAGAGAGCTGGCCAGGCGGCTGGTTACTGAACTGCAGCATT 960
 QY 961 TGNATTTGGGAGCACAATGCGAGAGAGAGTTGGAGCTGTCTCTGAGGCTAGTACC 1020
 DB 961 TGNATTTGGGAGCACAATGCGAGAGAGAGTTGGAGCTGTCTCTGAGGCTAGTACC 1020
 QY 1021 TGGGGCAAAAGTGGCGGTGAGAGAAATTCAGGTGTGATGTGACAGCCAGGCTTCTG 1080
 DB 1021 TGGGGCAAAAGTGGCGGTGAGAGAAATTCAGGTGTGATGTGACAGCCAGGCTTCTG 1080
 QY 1081 ACAGAGAGCGCGGTATGACGAGAGATGCAACAGAGAACTGTGCTCTCTGAGCGTGA 1140
 DB 1081 ACAGAGAGCGCGGTATGACGAGAGATGCAACAGAGAACTGTGCTCTCTGAGCGTGA 1140
 QY 1141 TCTTGACACACAGGCGCTGTACTATTGTCTGAGACACGTGAGCAAGGCAACCTGGTGA 1200
 DB 1141 TCTTGACACACAGGCGCTGTACTATTGTCTGAGACACGTGAGCAAGGCAACCTGGTGA 1200
 QY 1201 TTTGCGGACCCGGGGGTGAGAGCCCTCGTGAACACCGGCTCAAGCTCTGAGTTTCTG 1260
 DB 1201 TTTGCGGACCCGGGGGTGAGAGCCCTCGTGAACACCGGCTCAAGCTCTGAGTTTCTG 1260
 QY 1261 ACGTGCGCAGAGGCGATGAGTACCTGAGAGCAAGAACTTGTGACACCGGACCTGGCC 1320
 DB 1261 ACGTGCGCAGAGGCGATGAGTACCTGAGAGCAAGAACTTGTGACACCGGACCTGGCC 1320
 QY 1321 CCGGCAACATCTCTGCTCAAGAGACCTGTGCGCAAGAGTGTGAGGCTGTGCTCT 1380
 DB 1321 CCGGCAACATCTCTGCTCAAGAGACCTGTGCGCAAGAGTGTGAGGCTGTGCTCT 1380
 QY 1381 AAGCGAGCGGAAAGGGCTGAGCTCAAGCCGCTGCGCTCAAGTGAAGGCGCCGAGG 1440
 DB 1381 AAGCGAGCGGAAAGGGCTGAGCTCAAGCCGCTGCGCTCAAGTGAAGGCGCCGAGG 1440
 QY 1441 CTCTCAAAACACGGGAGTTTCAACGACGAGTGTGAGTGTGAGTGTGAGTGTGCTCT 1500
 DB 1441 CTCTCAAAACACGGGAGTTTCAACGACGAGTGTGAGTGTGAGTGTGAGTGTGCTCT 1500
 QY 1501 GGGAGGTCTTCTCATATGAGAGGGCTGCGTACCTTAATGTCTCTGAAAGAGTGTGCG 1560
 DB 1501 GGGAGGTCTTCTCATATGAGAGGGCTGCGTACCTTAATGTCTCTGAAAGAGTGTGCG 1560
 QY 1561 AGGCGCTGAGAAAGGGGTACCGCATGGAACCCCGAGGCGTGTCCAGGCGCCGTCGACG 1620
 DB 1561 AGGCGCTGAGAAAGGGGTACCGCATGGAACCCCGAGGCGTGTCCAGGCGCCGTCGACG 1620
 QY 1621 TCTCATATGAGCACTGCTGAGAGGAGAGCCCGCGCGCCGCAACCTTCCGCAAACTGG 1680
 DB 1621 TCTCATATGAGCACTGCTGAGAGGAGAGCCCGCGCGCCGCAACCTTCCGCAAACTGG 1680
 QY 1681 CCGAGAGCTGGCGCGGAGCTAGAGCATGAGTGTGCGCCAGCTCTGCTCAAGGAGAG 1740
 DB 1681 CCGAGAGCTGGCGCGGAGCTAGAGCATGAGTGTGCGCCAGCTCTGCTCAAGGAGAG 1740
 QY 1741 ACAGCGAGGCTCCACCTTCGCGCGAGAGCCCTGAGCCCAACCGGTGGGGGCTT 1800
 DB 1741 ACAGCGAGGCTCCACCTTCGCGCGAGAGCCCTGAGCCCAACCGGTGGGGGCTT 1800
 QY 1801 TGGCGCCAGAGAGACCGAGAGAGTGTGAGGCGGTGGGGGCACTGACAGAGGCCAAG 1860
 DB 1801 TGGCGCCAGAGAGACCGAGAGAGTGTGAGGCGGTGGGGGCACTGACAGAGGCCAAG 1860
 QY 1861 AGGCTTCAGGCGGCGAGTATCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
 DB 1861 AGGCTTCAGGCGGCGAGTATCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
 QY 1921 GGGCTTGGGCGGCGCGTGGAGACCCAGAGCTGTGAGAGATGATGATGATGATGATGATG 1980
 DB 1921 GGGCTTGGGCGGCGCGTGGAGACCCAGAGCTGTGAGAGATGATGATGATGATGATGATG 1980
 QY 1981 ATTCTAAGACTTAATAAAA 2000

DB 1981 ATTCTAAGACTTAATAAAA 2000
 RESULT 4
 US-08-876-882-1
 : Sequence 1, Application US/08876882
 : Patent No. 5981201
 : GENERAL INFORMATION:
 : APPLICANT: Avraham, Hava
 : APPLICANT: Groopman, Jerome E.
 : TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT
 : TITLE OF INVENTION: OF BREAST CANCER
 : NUMBER OF SEQUENCES: 9
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Hamilton, Brook, Smith & Reynolds P.C.
 : STREET: Two Millitia Drive
 : CITY: Lexington
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02173-4799
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: Windows
 : SOFTWARE: FastSeq for Windows Version 2.0b
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/876,882
 : FILING DATE: 16-JUN-1997
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/035,228
 : FILING DATE: 08-JAN-1997
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Doreen, Hogle M
 : REGISTRATION NUMBER: 36,361
 : REFERENCE/DOCKET NUMBER: NEDH97-01pa
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 781-861-6240
 : TELEFAX: 781-861-9540
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1987 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA
 : US-08-876-882-1
 Query Match 95.5%; Score 1909.8; DB 2; Length 1987;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1973; Conservative 0; Mismatches 7; Indels 9; Gaps 5;
 QY 1 CTCGCTTCCAAAGTTGTGAGCGCGGAGCCCTCGGGGTGTGACAGCCGCGCTGCGGAGGCC 60
 DB 8 CTCGCTTCCAAAGTTGTGAGCGCGGAGCCCTCGGGGTGTGACAGCCGCGCTGCGGAGGCC 67
 QY 61 TCTTGGGGGGGGGCGCGGGCGGCTGCGGGGGCGCCCTGAGCAGAAACAGAGAAACC 120
 DB 68 TCTTGGGGGGGGGCGCGGGCGGCTGCGGGGGCGCCCTGAGCAGAAACAGAGAAACC 127
 QY 121 AGGCTGGTCCAGTGGGAGCAGCTGCTCTGCTGTGCGAGCGCGCTGGCGGTGGCA 180
 DB 128 AGGCTGGTCCAGTGGGAGCAGCTGCTCTGCTGTGCGAGCGCGCTGGCGGTGGCA 187
 QY 181 GGCATTTCCAGAGGTCCCGCAGCTGTGACCACTTGTCTGAGTGTGCTCTGACCTGCTCAG 240
 DB 188 GGCATTTCCAGAGGTCCCGCAGCTGTGACCACTTGTCTGAGTGTGCTCTGACCTGCTCAG 247
 QY 241 TTTCCCTCTGGGGGCGATGCGGGGCGAGAGCTCTGCTGCTTCTCTGCGGGCATTTACG 300
 DB 248 TTTCC--TCNCGGGGGGATGCGGGGCGAGAGCTCTGCTGCTTCTCTGCGGGCATTTACG 305

QY 301 GCTGTGATTTCTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCGAGCCCTGGCACC 360
 DB 306 GCTGTGATTTCTGTGAGAACTTCCCGGGGTGAGCCCGCTTCTCGAGCCCTGGCACC 365
 QY 361 CCCCTCCGCTCAGCAGGATGCCAAGAGGGCGTGGGGCCCGGGCACCAGGTATATA 420
 DB 366 CCCCTCCGCTCAGCAGGATGCCAAGAGGGCGTGGGGCCCGGGCACCAGGTATATA 425
 QY 421 CCAAAATGCGAGCACCACCCCGGAGGAGGAGTGGCTTCCGCAAGGGGAGAGTGG 480
 DB 426 CCAAAATGCGAGCACCACCCCGGAGGAGTGGCTTCCGCAAGGGGAGAGTGG 485
 QY 481 TCACCACTCTGAGAGCCTCGAGAACAGAGTGGTACCGCGTCAAGACACACAGTGG 540
 DB 486 TCACCACTCTGAGAGCCTCGAGAACAGAGTGGTACCGCGTCAAGACACACAGTGG 545
 QY 541 GACAGAGAGGGGCTGCTGGAGGCTGGGGGCGCTGGGGAGGGGAGGGGCGCTCCGCGAGAC 600
 DB 546 GACAGAGAGGGGCTGCTGGAGGCTGGGGGCGCTGGGGAGGGGAGGGGCGCTCCGCGAGAC 605
 QY 601 CCAAGCTCAGCCTCATGCGCGTGGTTCACAGGGAGATCTCGGGCCAGAGAGGCTGTCCAC 660
 DB 606 CCAAGCTCAGCCTCATGCGCGTGGTTCACAGGGAGATCTCGGGCCAGAGAGGCTGTCCAC 665
 QY 661 AGCTGAGCCTCCGAGAGATGGGCTGTCTCTGTCGGGAGTCCGCGCCACACCCGGCG 720
 DB 666 AGCTGAGCCTCCGAGAGATGGGCTGTCTCTGTCGGGAGTCCGCGCCACACCCGGCG 725
 QY 721 ACTAGCTCTGTGCGTGGAGCTTTGGCGCGGAGCTCATCCGCGTGTGACACCGG 780
 DB 726 ACTAGCTCTGTGCGTGGAGCTTTGGCGCGGAGCTCATCCGCGTGTGACACCGG 785
 QY 781 AGGGCCACTCTCAATTCGATGAGCGCGTGTCTTTCGCAACCTCATGACATGATGGAGAC 840
 DB 786 AGGGCCACTCTCAATTCGATGAGCGCGTGTCTTTCGCAACCTCATGACATGATGGAGAC 845
 QY 841 ATTACAGCAAGAGCAAGGGCGCTATCTGTGACCAAGCTGGTGGAGCAAAAGCGGAAACAG 900
 DB 846 ATTACAGCAAGAGCAAGGGCGCTATCTGTGACCAAGCTGGTGGAGCAAAAGCGGAAACAG 905
 QY 901 GGACCAAGTGGCGCGGAGAGAGTGGCGGAGGGGCGTGTACTGAACCTGACAGATT 960
 DB 906 GGACCAAGTGGCGCGGAGAGAGTGGCGGAGGGGCGTGTACTGAACCTGACAGATT 965
 QY 961 TGACATTGGGAGCAGATCGAGAGGAGAGTGGAGCTGTCTCTGCAAGGGTGAATACC 1020
 DB 966 TGACATTGGGAGCAGATCGAGAGGAGAGTGGAGCTGTCTCTGCAAGGGTGAATACC 1025
 QY 1021 TGGGGCAAAAGGTGGCGCGTGAAGATATCAAGTGTGATGTGACAGCCGAGGCTTCTGG 1080
 DB 1026 TGGGGCAAAAGGTGGCGCGTGAAGATATCAAGTGTGATGTGACAGCCGAGGCTTCTGG 1085
 QY 1081 AGGAGAGCGCGCTCATGAGAGATGCAACAGAGAACTGGTGGCTCTCTGGGGGTGA 1140
 DB 1086 AGGAGAGCGCGCTCATGAGAGATGCAACAGAGAACTGGTGGCTCTCTGGGGGTGA 1145
 QY 1141 TCTTGACACAGGGGCTGTACATTTGATGAGACAGTGAAGGCAAGGGCAAGCTGTGAAT 1200
 DB 1146 TCTTGACACAGGGGCTGTACATTTGATGAGACAGTGAAGGCAAGGGCAAGCTGTGAAT 1205
 QY 1201 TTTCTGGGAGCGGGGGTGGAGCGCTGCAACACCGGCTAGTCTCTGCAAGTTTCTGTC 1260
 DB 1206 TTTCTGGGAGCGGGGGTGGAGCGCTGCAACACCGGCTAGTCTCTGCAAGTTTCTGTC 1265
 QY 1261 AGGTGGCCAGGAGCATGAGTACTGAGAGAGCAAGAGTGTGACACCGGAGCTGGCCG 1320
 DB 1266 AGGTGGCCAGGAGCATGAGTACTGAGAGAGCAAGAGTGTGACACCGGAGCTGGCCG 1325
 QY 1321 CCCGCAACATCTGTGCTCAGAGAGCTGGTGGCCAGAGTCAAGCACTTTGGCTGGCCA 1380
 DB 1326 CCCGCAACATCTGTGCTCAGAGAGCTGGTGGCCAGAGTCAAGCACTTTGGCTGGCCA 1385
 QY 1381 AAGCGAGAGGAAAGGGGCTAGACTCAAGCGGCTGCCCTCAAGTGGAGCGGCCCGGAGG 1440

DB 1386 AAGCGAGAGGAAAGGGGCTAGACTCAAGCGGCTGCCCTCAAGTGGAGCGGCCGAGG 1445
 QY 1441 CTCTCAACACGGAAGTTCCACAGCAAGTCCGAGTGTGAGTTTGGGGTGTGCTCT 1500
 DB 1446 CTCTCAACACG--GTTCCACAGCAAGTCCGAGTGTGAGTTTGGGGTGTGCTCT 1502
 QY 1501 GGGAGTCTTCTCATATGAGAGCGGCTCCGTACCCATAAATGTACTGAAAGAGTGTGCG 1560
 DB 1503 GGGAGTCTTCTCATATGAGAGCGGCTCCGTACCCATAAATGTACTGAAAGAGTGTGCG 1562
 QY 1561 AGGCGTGGAGAGAGGGGTACCGCATGGAACCCCGAGGGCTGTGAGGCCCGCTGACAG 1620
 DB 1563 AGGCGTGGAGAGAGGGGTACCGCATGGAACCCCGAGGGCTGTGAGGCCCGCTGACAG 1622
 QY 1621 TCTCATGAGAGCTGTGGAGAGCGAGCCCGCGCGGACACCTTCCGCAAACTGG 1680
 DB 1623 TCTCATGAGAGCTGTGGAGAGCGAG--CGGCCGCGCGGACACCTTCCGCAAACTGG 1681
 QY 1681 CCGAGAGCTGGCCCGGAGCTACAGCATGACAGTGGCCCAAGCTCTCAGGGCAGG 1740
 DB 1682 CCGAGAGCTGGCCCGGAGCTACAGCATGACAGTGGCCCAAGCTCTCAGGGCAGG 1741
 QY 1741 AGCGGAGCGGCTCAGACTGCGCCGAGAGAGGCTGTGACCCAGCGGTGGGGCT 1800
 DB 1742 AGCGGAGCGG--TCCACCTTGGCCCGAAGCAGAGGCTGTGACCCAGCGGT--GGCCT 1798
 QY 1801 TGGCCCAAGAGAGCAGAGATGAGAGTGGGCTGGGGGCTGACAGCAGGCCCAAG 1860
 DB 1799 TGGCCCAAGAGAGCAGAGATGAGAGTGGGCTGGGGGCTGACAGCAGGCCCAAG 1858
 QY 1861 AGGCTCAGGCGGGCAAGTCACTCTCTGTTGCCCAAGAGAGGCTGGCCAGCTAGG 1920
 DB 1859 AGGCTCAGGCGGGCAAGTCACTCTCTGTTGCCCAAGAGAGGCTGGCCAGCTAGG 1918
 QY 1921 GGCTCTGGGCGGCGGCTGAGACCCCAAGCTGCGCAAGATGATGCGCCGATTAAGACG 1980
 DB 1919 GGCTCTGGGCGGCGGCTGAGACCCCAAGCTGCGCAAGATGATGCGCCGATTAAGACG 1978
 QY 1981 ATTCTAAGG 1989
 DB 1979 ATTCTAAGG 1987

RESULT 5
 US-09-315-928-1
 ; Sequence 1, Application US/09315928
 ; Patent No. 6368796
 ; GENERAL INFORMATION:
 ; APPLICANT: Avraham, Hava
 ; APPLICANT: Groopman, Jerome E.
 ; TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF
 ; FILE REFERENCE: BREAST CANCER
 ; CURRENT APPLICATION NUMBER: US/09/315,928
 ; CURRENT FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: US 08/876,882
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: US 60/035,228
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 1987
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (263)...(1846)
 US-09-315-928-1

Query Match 95.5%; Score 1909.8; DB 4; Length 1987;
 Best Local Similarity 99.2%; Pred. No. 0;

Matches 1973; Conservative 0; Mismatches 7; Indels 9; Gaps 5;

OY	1	TTCCGCTCCAAATTGGCGACCGCGGACCCGCTTCGGGGTGTGCAGCCGGTCCGCGAGGCC	60
Db	8	CTCCGCTCCAAATTGGCGACCGCGGACCCGCTTCGGGGTGTGCAGCCGGTCCGCGAGGCC	67
OY	61	TCCTGGGGGGGGGGGGGGGGTCCGGGGGGGGCCCTTCGAGCAAGAAACAGGAAGAAC	120
Db	68	TCCTGGGGGGGGGGGGGGGGTCCGGGGGGGGCCCTTCGAGCAAGAAACAGGAAGAAC	127
OY	121	AGGCTCGGTCAGTGTGCAACCCAGTCCCTACCTCCTGTGCCAGCCGCGCTGTGTGCA	180
Db	128	AGGCTCGGTCAGTGTGCAACCCAGTCCCTACCTCCTGTGCCAGCCGCGCTGTGTGCA	187
OY	181	GGCCATTCCAGCGCGCCCGACTGTGACACTTGTGCTGATGTGGCTCTCACTCGCTAG	240
Db	188	GGCCATTCCAGCGCGCCCGACTGTGACACTTGTGCTGATGTGGCTCTCACTCGCTAG	247
OY	241	TTTCCCTTCGGGGGGGGCGATGGCGGGGGCGAGGCTCTGTGTTTCTGTGCGGGCATTTACG	300
Db	248	TTTCCCTTCGGGGGGGGCGATGGCGGGGGCGAGGCTCTGTGTTTCTGTGCGGGCATTTACG	305
OY	301	GCTGTGATCTTGTGTGAGCACTTCCCGGGGTGAGCCCGCCGTTCTCCGAGCTTGTGCAC	360
Db	306	GCTGTGATCTTGTGTGAGCACTTCCCGGGGTGAGCCCGCCGTTCTCCGAGCTTGTGCAC	365
OY	361	CCCGTCCCGTTCAGCGACAGATGGCCACAGAGGCGTGGGGCCCGGGACCCAGTGTATCA	420
Db	366	CCCGTCCCGTTCAGCGACAGATGGCCACAGAGGCGTGGGGCCCGGGACCCAGTGTATCA	425
OY	421	CCAAATGTGAGACACCCCGCCCAAGCCAGGGGAGCTGTGCTTCGCAAGGGCGAGCTGG	480
Db	426	CCAAATGTGAGACACCCCGCCCAAGCCAGGGGAGCTGTGCTTCGCAAGGGCGAGCTGG	485
OY	481	TCACCATCTGTGAGGGCTGTGAGAAACAAGACTGTGTACCGGCTCAAGCAACACACATG	540
Db	486	TCACCATCTGTGAGGGCTGTGAGAAACAAGACTGTGTACCGGCTCAAGCAACACACATG	545
OY	541	GACAGAGAGGGCTGTGTGACGCTGGGGCGCTGCGGGAGCGGAGAGCCCTTCGCGAGAC	600
Db	546	GACAGAGAGGGCTGTGTGACGCTGGGGCGCTGCGGGAGCGGAGAGCCCTTCGCGAGAC	605
OY	601	CCAAGCTCAGGCTATATGCCGCTGTTCACAGGGGAAGATTCGCGGGCAGGAGGCTGTCCAG	660
Db	606	CCAAGCTCAGGCTATATGCCGCTGTTCACAGGGGAAGATTCGCGGGCAGGAGGCTGTCCAG	665
OY	661	AGCTGCAAGCTTCGCGAGATGGGCTGTCTCTGTGTGCGGGAATTCGCGGGCACCCGGCG	720
Db	666	AGCTGCAAGCTTCGCGAGATGGGCTGTCTCTGTGTGCGGGAATTCGCGGGCACCCGGCG	725
OY	721	ACTTACGTCTGTGGTGAAGCTTTTGGCGGCAAGCTATCCACTTACCGCGCTGTGCACGGG	780
Db	726	ACTTACGTCTGTGGTGAAGCTTTTGGCGGCAAGCTATCCACTTACCGCGCTGTGCACGGG	785
OY	781	ACGCGCACTCACAATCATGATGAGGCGCGTTCCTTGTGCAACCTATGAGCAATGGGTGAGC	840
Db	786	ACGCGCACTCACAATCATGATGAGGCGCGTTCCTTGTGCAACCTATGAGCAATGGGTGAGC	845
OY	841	ATTACAGCAAGACAAAGGCGCTATCTTGCACCAAGCTGTGTGAGACCAAAAGCGGAACACG	900
Db	846	ATTACAGCAAGACAAAGGCGCTATCTTGCACCAAGCTGTGTGAGACCAAAAGCGGAACACG	905
OY	901	GGACCAAGTGTGGCGAGGAGGAGTGGCCAGGGGGGGTGGTACTGAACTTGACGATTT	960
Db	906	GGACCAAGTGTGGCGAGGAGGAGTGGCCAGGGGGGGTGGTACTGAACTTGACGATTT	965
OY	961	TGACATTGGGAGACAGATCGAGAGAGGAGATTGTGAGCTGTCTCTGAGGGTGAATACC	1020
Db	966	TGACATTGGGAGACAGATCGAGAGAGGAGATTGTGAGCTGTCTCTGAGGGTGAATACC	1025
OY	1021	TGGGGCAAAAGTGGCGCTGAAGAAATACAGTGTGATGTACAGCCAGGCTTTCCTGG	1080
Db	1026	TGGGGCAAAAGTGGCGCTGAAGAAATACAGTGTGATGTACAGCCAGGCTTTCCTGG	1085

QY	1081	ACGAGACGCGCCGTATAGCAAGTAATGCAACACGAGAACTGTGTGCTCTCTGGGCGCTGA	1140
Db	1086	ACGAGACGCGCGGTATGACGAAGATGGAACCGAGAACTGTGTGCTCTCTGGGCGCTGA	1145
QY	1141	TCCTGACACGAGGGGCTGTACATTTCTCATGTGAGCAGTGTAGCAAGGGCAACCTGGTGAAC	1200
Db	1146	TCCTGTACACGAGGGGCTGTACATTTCTCATGTGAGCAGTGTAGCAAGGGCAACCTGGTGAAC	1205
QY	1201	TTTCTGCGACCCGGGGGTGTGAGCCCTCGTGAACACCGCTCAGCTCTCTCAAGTTTTCTGTGC	1260
Db	1206	TTTCTGCGGACCCGGGGGTGTGAGCCCTCGTGAACACCGCTCAGCTCTCTCAAGTTTTCTGTGC	1265
QY	1261	ACGTGTGCGGAGGGCATGTAGTACTGTGGAGAGCAAGAACTTGTGCACCAGGACTGGCCG	1320
Db	1266	ACGTGTGCGGAGGGCATGTAGTACTGTGGAGAGCAAGAACTTGTGCACCAGGACTGGCCG	1325
QY	1321	CCCGCAACATCTCTGTCTCAGAGGACCTGGTGGGCCAAGGTATGAGGACTTTTGGCTGGCCA	1380
Db	1326	CCCGCAACATCTCTGTCTCAGAGGACCTGGTGGGCCAAGGTATGAGGACTTTTGGCTGGCCA	1385
QY	1381	AAGCCGAGCGGAGGGGGCTAGACATCAAGCCGGCTGCCGTCMAGTGAGAGCGGCCGAGG	1440
Db	1386	AAGCCGAGCGGAGGGGGCTAGACATCAAGCCGGCTGCCGTCMAGTGAGAGCGGCCGAGG	1445
QY	1441	CTCTCAAAACAGGGGAAGTTCACACGCAAGTGGATGTCTGGAGTTTGGGGTGTCTCT	1500
Db	1446	CTCTCAAAACAGGG--GTTCAACACGCAAGTGGATGTCTGGAGTTTGGGGTGTCTCT	1502
QY	1501	GGAGAGCTTCTCATATATGAGCGGGCTCCGCTCAAAATTCACCTGAAGAGGTGTCCG	1560
Db	1503	GGAGAGCTTCTCATATATGAGCGGGCTCCGCTCAAAATATTCACCTGAAGAGGTGTCCG	1562
QY	1561	AGGCCGTGAGAGGGGTATCCGCATGGAACCCCGAGGGCTGTCCAAGGCCCGCTGCAGC	1620
Db	1563	AGGCCGTGAGAGGGGTATCCGCATGGAACCCCGAGGGCTGTCCAAGGCCCGCTGCAGC	1622
QY	1621	TCCTCATATGAGAGTGTCTGGGAGGACAGACCCCGCGCCACACCTTCGCGAAATCG	1680
Db	1623	TCCTCATATGAGAGTGTCTGGGAGGACAGAG--CCGCGCCCGCGCCACACCTTCGCGAAATCG	1681
QY	1681	CCGGAAGACTGGCCCGGGAGCTTAGCAGTGCAGGTGGCCCGACCCCTCGGTCCAGGGCAG	1740
Db	1682	CCGGAAGACTGGCCCGGGAGCTTAGCAGTGCAGGTGGCCCGACCCCTCGGTTCAGGGCAG	1741
QY	1741	ACGCCGAGGGCTCCACCTCGGCCCGGAGCCAGGACGAGGCCCTGACCCACCCGGTGGGGCT	1800
Db	1742	ACGCCGAGAGG--TCCACCTCGGCCCGGAGCCAGGAGCCCTGACCCACCCGGT--GGCGCT	1798
QY	1801	TGCGCCCAAGAGGACCGAGAGAGTGGAGAGTGGCGGCTGGGGGCACTGACCAAGGCCCAAG	1860
Db	1799	TGCGCCCAAGAGGACCGAGAGAGTGGAGAGTGGCGGCTGGGGGCACTGACCAAGGCCCAAG	1858
QY	1861	AGGGTCCAGGGGGGCAAGTCACTCTCTCGTGGTCCCAACAGCAGGGGGCTGGGCCACGTAGG	1920
Db	1859	AGGGTCCAGGGGGGCAAGTCACTCTCTCGTGGTCCCAACAGCAGGGGGCTGGGCCACGTAGG	1918
QY	1921	GGCTCTGGGGCGCCGTGTGACACCCCAAGACCTTGGGAAGATGATCGGCCGATAAGACGG	1980
Db	1919	GGCTCTGGGGCGCCGTGTGACACCCCAAGACCTTGGGAAGATGATCGGCCGATAAGACGG	1978
QY	1981	ATTCTAAG 1989	
Db	1979	ATTCTAAG 1987	

RESULT 6
US-08-604-989A-11
: Sequence 11. Application US/08604989A
: Patent No. 5834208
: GENERAL INFORMATION:
: APPLICANT: Sakano, S.
: TITLE OF INVENTION: No. 5834208e1 Tyrosine Kinase

|||||
Db 1560 CCCCGTACAGCTCTCATGAGACAGCTGTGGAGAGAGCCCGCCGCGCACCTT 1619
QY 1670 CCGCAACTGGCCGAGAGAGCTGGCCCGGAGACTACGAGTACAGTGTGCTGCTGCTG 1729
Db 1620 CCGCAACTGGCCGAGAGAGCTGGCCCGGAGACTACGAGTACAGTGTGCTGCTGCTG 1679
QY 1730 CTCAGGAGAGAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1789
Db 1680 CTCAGGAGAGAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1739
QY 1790 GGTGGGCTTGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1849
Db 1740 GGT--GGCCCTTGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1797
QY 1850 CAGCCCAAG 1909
Db 1798 CAGCCCAAG 1857
QY 1910 CCGCAGTAG 1969
Db 1858 CCGCAGTAG 1917
QY 1970 GATAAGAGAGAGATCTTAAGAGACTCT 1994
Db 1918 GATAAGAGAGATCTTAAGAGACTCT 1942

RESULT 7
US-08-604-989A-10
Sequence 10, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-10
Query Match 76.0%; Score 1519.4; DB 2; Length 1521;

Best Local Similarity 99.9%; Pred. No. 9.5e-300;
Matches 1520; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 258 ATGCGGGGCGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 317
Db 1 ATGCGGGGCGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 318 GAATTCCTCCGGGGAG 377
Db 61 GAATTCCTCCGGGGAG 120
QY 378 AGATGCGCAACAG 437
Db 121 AGATGCGCAACAG 180
QY 438 CGCCCAAG 497
Db 181 CGCCCAAG 240
QY 498 TGGGAGAAACAG 557
Db 241 TGGGAGAAACAG 300
QY 558 GCAGCTGGGGGCTGCGGGAG 617
Db 301 GCAGCTGGGGGCTGCGGGAG 360
QY 618 CCGTGCTTCAAG 677
Db 361 CCGTGCTTCAAG 420
QY 678 GATGGGCTGTCGAG 737
Db 421 GATGGGCTGTCGAG 480
QY 738 AGCTTTGGCCGCGAG 797
Db 481 AGCTTTGGCCGCGAG 540
QY 798 GATGAGGCGCTGCTTCTGCAACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 857
Db 541 GATGAGGCGCTGCTTCTGCAACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 858 GGGCCTATCTGACCAAG 917
Db 601 GGGCCTATCTGACCAAG 660
QY 918 GAG 977
Db 661 GAG 720
QY 978 ATCGGAG 1037
Db 721 ATCGGAG 780
QY 1038 GTGAAGATATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1097
Db 781 GTGAAGATATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 1098 ACGAAGATGCAACAG 1157
Db 841 ACGAAGATGCAACAG 900
QY 1158 TACATTTGTCATGAG 1217
Db 901 TACATTTGTCATGAG 960
QY 1218 CGAGCCCTCGTGAACAG 1277
Db 961 CGAGCCCTCGTGAACAG 1020
QY 1278 GAGTACCTGAG 1337
Db 1278 GAGTACCTGAG

Db 1021 GAGTACCTGAGACGAAAGCTTTGTGACCGGACCTGGCCGCCGCAACATCTGTGTC 1080
 QY 1338 TCAGAGGACCTGTGGCCAGAGTCAAGCTTTGGCTGGCCAAAGCCGAGGAGGG 1397
 Db 1081 TCAGAGGACCTGTGGCCAGAGTCAAGCTTTGGCTGGCCAAAGCCGAGGAGGG 1140
 QY 1398 CTGAGCTCAAGCCGCTGCCCTGCAAGTGAGAGCGCCGAGGCTCTCAAAACGAGAG 1457
 Db 1141 CTGAGCTCAAGCCGCTGCCCTGCAAGTGAGAGCGCCGAGGCTCTCAAAACGAGAG 1200
 QY 1458 TTACACGACGAGTGTGATGTTTGGGTTGTGCTGGGAGGCTCTCATAT 1517
 Db 1201 TTACACGACGAGTGTGATGTTTGGGTTGTGCTGGGAGGCTCTCATAT 1260
 QY 1518 GAGAGGAGCTCGTACCTAAATGTCACTAAAGAGGTGTGAGAGCCGTGAGAGAGG 1577
 Db 1261 GAGAGGAGCTCGTACCTAAATGTCACTAAAGAGGTGTGAGAGCCGTGAGAGAGG 1320
 QY 1578 TACCGATGGAACCCCCCGAGGCTGTCCAGGCCCCGTGACGTCTCATGACAGCTGC 1637
 Db 1321 TACCGATGGAACCCCCCGAGGCTGTCCAGGCCCCGTGACGTCTCATGACAGCTGC 1380
 QY 1638 TGGGAGGACAGCCCGCCGCGCCGACCCCTTCCGAAACTGGCCGAGAGTGGCCCG 1697
 Db 1381 TGGGAGGACAGCCCGCCGCGCCGACCCCTTCCGAAACTGGCCGAGAGTGGCCCG 1440
 QY 1698 GAGCTACGAGTGTGAGTGGCCGAGGCTCGTCTCAGGGGAGGAGCCGAGGCTCCAC 1757
 Db 1441 GAGCTACGAGTGTGAGTGGCCGAGGCTCGTCTCAGGGGAGGAGCCGAGGCTCCAC 1500
 QY 1758 TCGCCCCGAAAGCCAGAGGCC 1778
 Db 1501 TCGCCCCGAAAGCCAGAGGCC 1521

RESULT 8

US-08-604-989A-9
 ; Sequence 9, Application US/08604989A
 ; Patent No. 5834208
 ; GENERAL INFORMATION:
 ; APPLICANT: Sakano, S.
 ; TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/604,989A
 ; FILING DATE: February 23, 1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Charles E. Miller
 ; REGISTRATION NUMBER: 24,576
 ; REFERENCE/DOCKET NUMBER: 1920-026
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1398 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: human
 ; STRAIN: UT-7
 ; US-08-604-989A-9

Query Match 69.9%; Score 1398; DB 2; Length 1398;
 Best Local Similarity 100.0%; Pred. No. 4e-275;
 Matches 1398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 381 ATGCCAAGAGGCGGCGGCGCCCGGCGACCCAGTATATACAAAGGAGACACACCCG 440
 Db 1 ATGCCAAGAGGCGGCGGCGCCCGGCGACCCAGTATATACAAAGGAGACACACCCG 60
 QY 441 CCCAAGCCAGGAGGAGCTGCGCTTCCGAAAGGCGAGTGTATACATCTCGAGGCGTGC 500
 Db 61 CCCAAGCCAGGAGGAGCTGCGCTTCCGAAAGGCGAGTGTATACATCTCGAGGCGTGC 120
 QY 501 GAGACAAAGAGCTGTACCGCGTCAAGACACACACACAGTGTGACAGAGGAGGCTGTGCA 560
 Db 121 GAGACAAAGAGCTGTACCGCGTCAAGACACACACAGTGTGACAGAGGAGGCTGTGCA 180
 QY 561 GCTGGGGGCGTGGGGAGGGGAGGCGCTCTCCGAGACCCCAAGCTCAAGCTCATAGCCG 620
 Db 181 GCTGGGGGCGTGGGGAGGGGAGGCGCTCTCTCCGAGACCCCAAGCTCAAGCTCATAGCCG 240
 QY 621 TGGTTCACAGGGAAGATCTCGGGCCAGAGAGGTGTTCACAGCTGAGGCTCCGAGAGT 680
 Db 241 TGGTTCACAGGGAAGATCTCGGGCCAGAGAGGTGTTCACAGCTGAGGCTCCGAGAGT 300
 QY 681 GGGCTTCTCTGTGGTGGGAGTCCGCGCCACCCCGGAGTACGTCTGTGCTGTGAGC 740
 Db 301 GGGCTTCTCTGTGGTGGGAGTCCGCGCCACCCCGGAGTACGTCTGTGCTGTGAGC 360
 QY 741 TTTGGCCGGGAGGTATCACTACCGGCTGTGACAGGAGGCGCCACCTCAAAATCAT 800
 Db 361 TTTGGCCGGGAGGTATCACTACCGGCTGTGACAGGAGGCGCCACCTCAAAATCAT 420
 QY 801 GAGGCGGCTTCTTCTGCAACCTCATGAGCATGTGTGAGCATTTACAGCAAGAGAGGCG 860
 Db 421 GAGGCGGCTTCTTCTGCAACCTCATGAGCATGTGTGAGCATTTACAGCAAGAGAGGCG 480
 QY 861 GCTATCTGCACCAAGCTGTGTGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 920
 Db 481 GCTATCTGCACCAAGCTGTGTGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 QY 921 GAGCTGGCCAGGCGGCGGCTGTGAGTACCTGTGACATTTGACATTTGGGAGCACAGATC 980
 Db 541 GAGCTGGCCAGGCGGCGGCTGTGAGTACCTGTGACATTTGACATTTGGGAGCACAGATC 600
 QY 981 GGAAGAGGAGAGTTTGGAGCTGTCTGCAAGGAGTGTACCTGGGCAAAAGGTTGGCCGTG 1040
 Db 601 GGAAGAGGAGAGTTTGGAGCTGTCTGCAAGGAGTGTACCTGGGCAAAAGGTTGGCCGTG 660
 QY 1041 AAGATATCAAGTGTATGTACAGCCAGGCTTCTGTGACAGAGGCGCTCATGACG 1100
 Db 661 AAGATATCAAGTGTATGTACAGCCAGGCTTCTGTGACAGAGGCGCTCATGACG 720
 QY 1101 AAGATCAACAGAGAACTGTGTGCTGCGGGGCTGATCTGACACAGGAGGCTGTAC 1160
 Db 721 AAGATCAACAGAGAACTGTGTGCTGCGGGGCTGATCTGACACAGGAGGCTGTAC 780
 QY 1161 ATTGTATGAGACAGTGTGAGCAAGGCAACCTGTGTAATTTCTGCGAGCCCGGGGTGCA 1220
 Db 781 ATTGTATGAGACAGTGTGAGCAAGGCAACCTGTGTAATTTCTGCGAGCCCGGGGTGCA 840
 QY 1221 GCCCTGTGTAACAGCGCTAGTCTGTGCAAGTTTCTGTGACAGTGTGCGAGGAGATGAG 1280
 Db 841 GCCCTGTGTAACAGCGCTAGTCTGTGCAAGTTTCTGTGACAGTGTGCGAGGAGATGAG 900
 QY 1281 TACCTGGAGAGCAAGAGTGTGACAGCGGAGCTGTGCGGCGGCAACATCTGTGATCA 1340
 Db 901 TACCTGGAGAGCAAGAGTGTGACAGCGGAGCTGTGCGGCGGCAACATCTGTGATCA 960

QY	1341	GAGGACCTGGTGGCCAAAGTCATGGCACTTGGCTGGCCAAAGCCAGAGCGGAAGGGCTA	1400
Db	961	GAGGACCTGGTGGCCAAAGTCATGGCACTTGGCTGGCCAAAGCCAGAGCGGAAGGGCTA	1020
QY	1401	GACTCAAGCCGGCTGGCCGCTCAAGTCAGCGAGCGCCGAGGCTCTCAAAACAGGGGAATTC	1460
Db	1021	GACTCAAGCCGGCTGGCCGCTCAAGTCAGCGAGCGCCGAGGCTCTCAAAACAGGGGAATTC	1080
QY	1461	ACCAAGCAAGTCGGATGTCGTGGAGTTTGGGTGCTGCTCTGGGAAGTCTTCTCATPATGA	1520
Db	1081	ACCAAGCAAGTCGGATGTCGTGGAGTTTGGGTGCTGCTCTGGGAAGTCTTCTCATPATGA	1140
QY	1521	CGGAGCTCCGTACCCCTAAATATCTCACTGAAAGAGTGTGGAGGCCGTGGAGAAAGGGGTAC	1580
Db	1141	CGGAGCTCCGTACCCCTAAATATCTCACTGAAAGAGTGTGGAGGCCGTGGAGAAAGGGGTAC	1200
QY	1581	CGCATGGAAACCCCGCAGAGGCTGTCCAGAGGCCCGCTGCACGTCTCATGAGCAGAGTGTGG	1640
Db	1201	CGCATGGAAACCCCGCAGAGGCTGTCCAGAGGCCCGCTGCACGTCTCATGAGCAGAGTGTGG	1260
QY	1641	GAGGCAGAGCCCGCGCGCGGCCCAACCTTCGCGAAATGGGCGGAGAGCTGGCCCGGGAG	1700
Db	1261	GAGGCAGAGCCCGCGCGGCCCAACCTTCGCGAAATGGGCGGAGAGCTGGCCCGGGAG	1320
QY	1701	CTACGAGATGAGGTGGCCCAAGCCTTCGCTCTCAAGGAGAGACAGCGACGGCTCCACCTCG	1760
Db	1321	CTACGAGATGAGGTGGCCCAAGCCTTCGCTCTCAAGGAGAGACAGCGACGGCTCCACCTCG	1380
QY	1761	CCCCGAAGCCAGAGAGCCC	1778
Db	1381	CCCCGAAGCCAGAGAGCCC	1398

```

RESULT 9
US-09-741-154-1
: Sequence 1, Application US/09741154
: Patent No. 6437110
: GENERAL INFORMATION:
: APPLICANT: BEASLEY, Ellen M. et al
: TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
: TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: CLO01061
: CURRENT APPLICATION NUMBER: US/09/741,154
: CURRENT FILING DATE: 2000-12-21
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 1713
: TYPE: DNA
: ORGANISM: Human,
US-09-741-154-1

```

Query Match	68.8%	Score 1377;	DB 4;	Length 1713;
Best Local Similarity	99.3%	Pred. No. 7.6e-271;		
Matches 1383; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

QY	608	CAGCCTCATGCCGCTGTTCCACGGGGAAGATTCCGGGCCAGGAGGCTGTCCAGCAGCTGCA	667
Db	297	CAGAGCTTCTGGTGGTTCCACGGGGAAGATTCCGGGCCAGGAGGCTGTCCAGCAGCTGCA	356
QY	668	GCTCTCCGAGGATGGGCTTCTCGTGTGCGGGAGATCCGGCGGCGACCCCGGCGACTTACGT	727
Db	357	GCCTCCGAGGATGGGCTTCTCTGTGTGCGGGAGATCCGGCGGCGACCCCGGCGACTTACGT	416
QY	728	CTTGTGCGTGAAGCTTTTGGCCGCGAGCTATCCACTTACCGCGTGTGCAACCGCGAGCGCA	787
Db	417	CTGTGCGTGAAGCTTTTGGCCGCGAGCTATCCACTTACCGCGTGTGCAACCGCGAGCGCA	476
QY	788	CTCTACAAATGATATAGGCGCGTGTCTTCTCTGCAACCTATGGACATGTGTGAAGCAATTACAG	847
Db	477	CTCTACAAATGATATAGGCGCGTGTCTTCTCTGCAACCTATGGACATGTGTGAAGCAATTACAG	536

QY	848	C A A G G A C A A G G G C C C T A T C T G C A C C A A G C T G T G A G C A A A G C G G A A C A A C G G G A C C A A	907
Db	537	C A A G G A C A A G G G C C C T A T C T G C A C C A A G C T G T G A G C A A A A G G G A A A C C G G A C C A A	596
QY	908	G T C G C C G A G A G A G A G C T G G G C C A G G G G G G C T G T T A C T G A A C C T G G A G C A T T T G A C A T T	967
Db	597	G T C G C C G A G A G A G A G C T G G C C A G G G G G G C T G T T A C T G A A C C T G C A C A C T T T G A C A T T	656
QY	968	G G G A G C A C A G A T C G G A G A G G A G A G A G T T T G G A G C T G T C C T G C A G G G T A G T A C T G G G G C A	1027
Db	657	G G G A G C A C A G A T C G G A G A G G A G A G A G T T T G G A G C T G T C C T G C A G G G T A G T A C T G G G G C A	716
QY	1028	A A A G G T G C C G T G A A G A A T A T C A G T G T G A T G T C A C A G C C A A G C C A G C C T T C C T G G A C G A G A C	1087
Db	717	A A A G G T G C C C T G A A G A A T A T A C A G T G T G A T G T A C A G C C A A G C C A G C C A T T C T G G A C G A G A C	776
QY	1088	G G C C G T A T G A C G A A G A T G C A A C A C A G A A C T T G T G C G T C T C C T G G G C G T G A T C T G C A	1147
Db	777	G G C C G T A T G A C G A A G A T G C A A C A C A G A A C C T G T G C G T C T C C T G G G C G T G A T C T G C A	836
QY	1148	C C A G G G C C T G T A C A T T T G T C A T G A G A C A C T G A C A A G G C A A C C T G G T G A A C T T T C T G C G	1207
Db	837	C C A G G G C C T G T A C A T T T G T C A T G A G A C A C T G A C A A G G C A A C C T G G T G A A C T T T C T G C G	896
QY	1208	G A C C C G G G G T G A C A C C C C C T G T G A A C A C C G C T C A C T C T G C A G T T T T C T G C A C A G T G G C	1267
Db	897	G A C C C G G G G T G A C A C C C C C T G T G A A C A C C G C T C A C T C T G A A G T T T T C T G C A C A G T G G C	956
QY	1268	C G A G G C A T G A G A G T A C T C G G A G A C A G A A G A A G C T T G T G C A C C G G A C C T G G C C G C C G A A	1327
Db	957	C G A G G C A T G A G A G T A C T C G G A G A C A G A A G A A G C T T G T G C A C C G C A C C T G G C C G C C G A A	1016
QY	1328	C A T C C T G T C T C A C A G A G A C C T G T G G G C C A A G A G T A G A G A C A T T T T G C C T G G G C C A A G C G A	1387
Db	1017	C A T C C T G T C T C A C A G A G A C C T G T G G G C C A A G A G T A G A G A C A T T T T G C C T G G G C C A A G C G A	1076
QY	1388	G C G G A A G G G G C T A G A C T C A A G C C G G C T G C C C G T C A A G T G A C A G G C G G C C G A G G C T C T C A A	1447
Db	1077	G C G G A A G G G G C T A G A C T C A A G C C G G G C T G C C C G T C A A G T G A C A G G C G G C C G C C G A G G C T C T C A A	1136
QY	1448	A C A G G G A A G T T C A C A G A C A A G T G G A A G T T G A G T T T T G G G T G G C T G C T G C T G G A G A G T	1507
Db	1137	A C A G G G A A G T T C A C A G A C A A G T G G A A G T T G A G T T T T G G G T G T G C T G T G G A G A G T	1196
QY	1508	C T T C T C A T A T G A G A G G G C T C C G T A C C C T A A A A T T C A C T G A A A G A G A G T G T C T G A G A G C C G T	1567
Db	1197	C T T C T C A T A T G A G A G G G C T C C G T A C C C T A A A A T T C A C T G A A A G A G A G T G T C T G A G A G C C G T	1256
QY	1568	G G A A A A G G G G T A C C G C A T G A A A C C C C C G A G A G C T G T C A G A G G C C C C T G C A C A G T C T C A T	1627
Db	1257	G G A A A A G G G G T A C C G C A T G A A A C C C C C G A G A G C T G T C A G A G G C C C C T G C A C A G T C T C A T	1316
QY	1628	G A G A G C T G C T G G A G A G C A G A C C C G C C G C C G C C A C A C C T T T G C G A A A C T G G C C G A G A A	1687
Db	1317	G A G A G C T G C T G G A G A G C A G A C C C C C G C C G C C G C C A C A C C T T T G C G A A A C T G G C C G A G A A	1376
QY	1688	G C T G G C C C G G A G C T A C G A C A G A G T G C C C A G C C T C C G T C T C A G A G G C A G A C G C G A	1747
Db	1377	G C T G G C C C C G G A G C T A C G A C A G A G T G C C C C A G C C T C C G T C T C A G A G G C A G A C G C G A	1436
QY	1748	C G G T C C A C C T C G C C C C G A A G C C A G A G A C C T G A C C C A C C C G G T G G G G G C C C T T G G C C C	1807
Db	1437	C G G T C C A C C T C G C C C C G A A G C C A G A G A C C T G A C C C A C C C G G T G G G G G C C C T T G G C C C	1496
QY	1808	A G A G A C C G A G A G T G G A G A G T C G G A C G T G G G G G C A C T G A C C A G G C C C A A G A G A G G T C C	1867
Db	1497	A G A G A C C G A G A G T G G A G A G T C G G A C G T G G G G G C A C T G A C C A G G C C C A A G A G A G G T C C	1556
QY	1868	A G G G G G G C A A G T A C T C T C T G T G T G C C C A C A G C A G A G G G G C T G G C C A C A G T A G A G G G C T C T G	1927
Db	1557	A G G G G G G C A A G T A C T C T C T G T G T G C C C A C A G C A G G G G C T G G C C C A C A G T A G A G G G G C T C T G	1616

QY 1928 GCGGCGCCGTGGACACCCCGACACTGCGAAGATGATCCGCCATAAAGACGATTTCTAA 1987
DB 1617 GCGGCGCCGTGGACACCCCGACACTGCGAAGATGATCCGCCATAAAGACGATTTCTAA 1676
QY 1988 GGAAGCTTAAAAA 2000
DB 1677 GGAAGCTTAAAAA 1689

RESULT 10
US-08-604-989A-8
Sequence 8, Application US/08604989A
Patent No. 5834208
GENERAL INFORMATION:
APPLICANT: Sakano, S.
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/604,989A
FILING DATE: February 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Charles E. Miller
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 1920-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: human
STRAIN: UT-7
US-08-604-989A-8

Query Match 36.9%; Score 738; DB 2; Length 738;
Best Local Similarity 100.0%; Pred. No. 3e-141;
Matches 738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 954 CAGCATTGACATTGGAGCAGACAGATCGAGAGGAGAGTGGAGCTGTCTGACGGGT 1013
DB 1 CAGCATTGACATTGGAGCAGACAGATCGAGAGGAGAGTGGAGCTGTCTGACGGGT 60
QY 1014 GAGTACCTGGGGCAAAAGTGGCCGTGAGAGATATCAAGTGTGATGACAGCCAGGCC 1073
DB 61 GAGTACCTGGGGCAAAAGTGGCCGTGAGAGATATCAAGTGTGATGACAGCCAGGCC 120
QY 1074 TTCTGGAGAGAGGCGGCTCATGAGAGAGATGCAACACAGAAACCTGGTCTCTCTG 1133
DB 121 TTCTGGAGAGAGGCGGCTCATGAGAGAGATGCAACACAGAAACCTGGTCTCTCTG 180
QY 1134 GGGGTATCTGACACAGAGGGGTGTACATGTGATGAGAGCAGTGAAGCAAGGCAACCTG 1193
DB 181 GGGGTATCTGACACAGAGGGGTGTACATGTGATGAGAGCAGTGAAGCAAGGCAACCTG 240

QY 1194 GTGAACCTTTTCGGGACCCCGGGGTGAGACCCCTGTGTAAACCCGCTCAGCTCTGCAAGTTT 1253
DB 241 GTGAACCTTTTCGGGACCCCGGGGTGAGACCCCTGTGTAAACCCGCTCAGCTCTGCAAGTTT 300
QY 1254 TCTGTGACGTGGCGCGAGGCGATGAGTACTGTGAGAGCAAGAAAGCTGTGACCCGGAC 1313
DB 301 TCTGTGACGTGGCGCGAGGCGATGAGTACTGTGAGAGCAAGAAAGCTGTGACCCGGAC 360
QY 1314 CTGGCCGCGCGCAACATCTGTGTCTCAGAGACCTGGTGGCCCAAGGTCAAGCTTTGGC 1373
DB 361 CTGGCCGCGCGCAACATCTGTGTCTCAGAGACCTGGTGGCCCAAGGTCAAGCTTTGGC 420
QY 1374 CTGGCCAAAGCCGAGCGGAAGGGGCTGACTCAAGCCGCTGCCCTCAAGTGAAGCGCG 1433
DB 421 CTGGCCAAAGCCGAGCGGAAGGGGCTGACTCAAGCCGCTGCCCTCAAGTGAAGCGCG 480
QY 1434 CCGGAGCTCTCAACACAGGGGAAGTTCACAGCAAGTGTGATGATTTTGGGTG 1493
DB 481 CCGGAGCTCTCAACACAGGGGAAGTTCACAGCAAGTGTGATGATTTTGGGTG 540
QY 1494 CTGCTGGAGAGCTCTCTCATATGAGACGGGCTCCGTAACCTTAAATGTCAAGAAAG 1553
DB 541 CTGCTGGAGAGCTCTCTCATATGAGACGGGCTCCGTAACCTTAAATGTCAAGAAAG 600
QY 1554 GTGTGGAGGCGGTGGAAGGGGTACCGCATGGAACCCCGAGGGCTGTCCAGGCGCC 1613
DB 601 GTGTGGAGGCGGTGGAAGGGGTACCGCATGGAACCCCGAGGGCTGTCCAGGCGCC 660
QY 1614 GTGACGCTCTCATGAGACACTGTGTGAGAGCAGAGCCCGCGGCGCAACCTTCCG 1673
DB 661 GTGACGCTCTCATGAGACACTGTGTGAGAGCAGAGCCCGCGGCGCAACCTTCCG 720
QY 1674 AAATGCGCGAGAAAGCTG 1691
DB 721 AAATGCGCGAGAAAGCTG 738

RESULT 11
US-09-741-154-3
Sequence 3, Application US/09741154
Patent No. 6437110
GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M. et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C1001061
CURRENT APPLICATION NUMBER: US/09/741,154
CURRENT FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 16389
TYPE: DNA
ORGANISM: Human
US-09-741-154-3

Query Match 22.8%; Score 455; DB 4; Length 16389;
Best Local Similarity 100.0%; Pred. No. 1.3e-83;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1541 GTCACTGAAAGAGGTGTGAGGCGGTGAGAAAGGGGTACCCGATGAAACCCCGAGGG 1600
DB 13941 GTCACTGAAAGAGGTGTGAGGCGGTGAGAAAGGGGTACCCGATGAAACCCCGAGGG 14000
QY 1601 CTGTCCAGGCGCGGTGAGAGTCTCTATGAGACACTGCTGGAGGAGAGAGCCCGCGCG 1660
DB 14001 CTGTCCAGGCGCGGTGAGAGTCTCTATGAGACACTGCTGGAGGAGAGAGCCCGCGCG 14060
QY 1661 GCCACCTTCGCAAACTGGCGAGAAAGCTGGCCCGGAGACTTACGAGATGACAGTCCCG 1720
DB 14061 GCCACCTTCGCAAACTGGCGAGAAAGCTGGCCCGGAGACTTACGAGATGACAGTCCCG 14120

TITLE: pp60c-src
JOURNAL: Molecular and Cellular Biology
VOLUME: 5
ISSUE: 5
PAGES: 1122-1129
DATE: May, 1985
PUBLICATION INFORMATION:
AUTHORS: Tanaka, Akio
AUTHORS: Gibbs, Carol P.
AUTHORS: Arthur, Richard R.
AUTHORS: Anderson, Stephen K.
AUTHORS: Kung, Hsing-Jien
AUTHORS: Fujita, Donald J.
TITLE: DNA Sequence Encoding the
TITLE: Amino-Terminal Region of the Human c-src
TITLE: Protein: Implications of Sequence
TITLE: Divergence among src-Type Kinase
TITLE: Oncogenes
JOURNAL: Molecular and Cellular Biology
VOLUME: 7
ISSUE: 5
PAGES: 1978-1983
DATE: May, 1987
US-07-820-011A-3

Query Match 10.6%; Score 212.4; DB 1; Length 1611;
Best Local Similarity 55.9%; Pred. No. 1.5e-34;
Matches 455; Conservative 0; Mismatches 341; Indels 18; Gaps 2;

QY 897 CACGGACCAAGTCGGCCGAGAGAGCTGGCCAGGCGCGCTGTTACTGAACCTGCAG 956
DB 745 CCCACCTCGAAGCCGAGACTCAGGCGCTGGCCAAAGATGCTCGGAGATCCCTCGGAG 804
QY 957 CATTGTACATTGGAGACACAGATCGAGAGAGAGATTGTGAGCTCTCGAGGGTGAG 1016
DB 805 TGGCTCGCGCTGGAGTCAAGCTGGCCAGGCGCTGTTGGCAGGTGAGATGGGAGC 864
QY 1017 TACCTGGGCAAAAGGTGGCCCTGAAGAAATATCAAGTGTGAG-----TGACAGCCAG 1070
DB 865 TGAACGGTACCAACCGAGGTGGCCATCAAAACCTGAACCTGGACAGATGTCACAGAG 924
QY 1071 GCCTTCCTGAGAGAGCGGCCCTGATCAGAGAGATGACACAGAGAACTGGTGGCTGC 1130
DB 925 GCCTTCCTGAGAGAGCGGCCCTGATCAGAGAGATGACAGAGAACTGGTGGCTGC 984
QY 1131 CTGGGCGGTATCTCTGACACAGAGGGGTGTACATTGTATGAGAGACCTGAGCAAGGCAAC 1190
DB 985 TATGCTGTGTTTCAAGAGAGAGCCCATTTATCTGTCAGAGATGATGAGCAAGGGAGT 1044
QY 1191 CTGGTGAATCTTCTGCGGACCCGGGGTGAAGCCCTCTGTGAACACCGCTAGCTCTGCAG 1250
DB 1045 TTGCTGTGACTTCTCAAGGGGAGAGACAGCAAGTACTCTCGGCTGCTCAGCTGGTGGAC 1104
QY 1251 TTTTCTCTGACGTGGCGAGGCGATGATGATCCTGAGAGAGAAAGTGTGGACCGGC 1310
DB 1105 ATGGCTGCTCAATGCTCTCAGAGCATGGCTGATGAGAGAGTGAATGATGATGATGATG 1164
QY 1311 GACCTGGCGCCCGCAACATCTGTCTCAGAGAGAGAGTGTGAGAGTGTGAGAGTGTG 1370
DB 1165 GACCTTCTGTGAGAGCAATCTGTGTGGAGAGAACTGTGTGTGCAAAAGTGGCGACTTT 1224
QY 1371 GGCCTGGCCAAAGCC-----GAGCGGAGAGGGGCTGTGACTCAAGCCGGCTGCC 1418
DB 1225 GGGCTGGCTCGGCTCTATTGAACAATGATGATGATGATGATGATGATGATGATGATG 1284
QY 1419 GTCAAGTGAAGCGCGCGAGGCTCAACAACGAGAAATGATGATGATGATGATGATGATG 1478
DB 1285 ATCAAGTGAAGCGCGCTCAACAACGAGAAATGATGATGATGATGATGATGATGATGATG 1344
QY 1479 TGAAGTGTGGGTGCTGTGAGAGTCTTCTCATGATGAGAGAGGCTGATGATGATGATG 1538
DB 1345 TGGTCTTGGGATCTGTGACTGATGATGATGATGATGATGATGATGATGATGATGATG 1404

QY 1539 ATGTCACTGAAGAGGTGTGCGAGCGCTGGAGAGAGGGGTACCGCATGGAACCCCGAG 1598
DB 1405 ATGTGAACCGGAGGTGTGCGAGCGCTGGAGAGAGGGGTACCGCATGGAACCCCGAG 1464
QY 1599 GCGTGTCAAGCGCCCTGACAGTCTCTCATGAGAGAGTGTGGAGAGCAAGCCCGCC 1658
DB 1465 GAGTGTCCGAGTCTCTGACAGCAAGCTCATGAGTGTGCGAGAGAGAGAGAGAGAG 1524
QY 1659 CGGCGACCTTCCGCAACTGGCGGAGAGAGTGG 1692
DB 1525 CGGCGACCTTCCGCAACTGTGAGAGAGAGTGG 1558

RESULT 14
PCT-US93-00445-3
Sequence 3, Application PC/TUS9300445
GENERAL INFORMATION:
APPLICANT: Bell, Leonard
APPLICANT: Madril, Joseph A.
APPLICANT: Warren, Stephen L.
APPLICANT: Luthinger, Daniel J.
TITLE OF INVENTION: Genetically Engineered
TITLE OF INVENTION: Endothelial Cells
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 760 Kb storage
COMPUTER: DELL 486/50
OPERATING SYSTEM: DOS 5.0
SOFTWARE: Displaywrite 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00445
FILING DATE: 19930105
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/820,011
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-101PCT
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
POSITION IN GENOME:
CHROMOSOME/SEGMENT: Chromosome 20
PUBLICATION INFORMATION:
AUTHORS: Anderson, Stephen K.
AUTHORS: Gibbs, Carol P.
AUTHORS: Tanaka, Akio
AUTHORS: Kung, Hsing-Jien
AUTHORS: Fujita, Donald J.
TITLE: Human Cellular src Gene:
TITLE: Nucleotide Sequence and Derived Amino
TITLE: Acid Sequence of the Region Coding for
TITLE: the Carboxy-Terminal Two-Thirds of

Query Match 10.0%; Score 199.2; DB 1; Length 1602;
Best Local Similarity 55.7%; Pred. No. 7.2e-32;
Matches 453; Conservative 0; Mismatches 343; Indels 18; Gaps 3;

```
OY 897 CACGGGACCAAGTCCGCCGAGAGAGAGTGGCCAGGCGGGCTGGTACTGACCTGCAG 956
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 736 CCCAGTCCAAAGCCCGAGACCCAGGAGCTCGCCAGAGGACGCGTGGAAATCCCGGGGAG 795
OY 957 CATTGACATGGGAGACAGATCGAGAGGAGATTGGAGCTGTCTCTGACGAGG---T 1013
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 796 TCGCTGCGGCTGGAGGTGAAGCTGGGCGAGGCTCTTTGGAGAGTCTGGATGGGGAGC 855
OY 1014 GAGTACTGGGGCAAAAGGCGCTGTAAGATATACAG---TGTGATGTGACAGCCGAG 1070
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 856 TGGAAAGGACACACCAAGAGTGGCCATAAAGACTGAAAGCCCGCAACATGTCCCGGAG 915
OY 1071 GCCTTCCTGGAGAGACGCGGCTCATGACGAAGATGCAACAGAGACCTGGTGGCTCTC 1130
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 916 GCCTTCCTGGAGAGACCCAGTGAAGAGCTCGGCATGAGAGCTGGTTGAGCTG 975
OY 1131 CTGGGCGTGAATCCTCACCAGGAGGCTGTACATTGTATGAGACAGCTGAGCAAGGGCAAC 1190
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 976 TACGCAAGTGTGTGCGAAGAGCCCATCTACATGCTACTGATGATGAGCAAGGGGAGC 1035
OY 1191 CTGGTGAACCTTCTGCGGACCCGCGGCTGAGCCCTGTAACACCCGCTCAGCTCTGCAG 1250
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1036 CTCCGTGATTTCTGTAAGGAGAGATGGGCAAGTACTGCGGCTCCACAGCTCGTCAT 1095
OY 1251 TTTTCTGCAAGTGGCGGAGGCGCATGAGTACTGAGAGAGCAAGAGCTTGTGACCCGC 1310
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1096 ATGGCTGCTCAAGATTGATCCGGCATGGCTATGTGAGAGGATGACTACGTCACCGA 1155
OY 1311 GACCTGAGCGCCCGGCAACATCTGAGTCTAGAGACCTGAGTGGCCAGGTACGACTTT 1370
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1156 GACCTGAGCGGCGGCGCAACATCTGAGTGGGGGAGAACTGTGTGCAAGTGGCTGACTTT 1215
OY 1371 GGCTGCGCAAAAGCCGAGCGGAGGGGCTAGACTCAAGCCGGC-----TGCCC 1418
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1216 GGGCTGGCACGCGCTCATGAGAGACAGAGTACACAGCAGGCAAGTGCCAAAGTTCCC 1275
OY 1419 GTCAAGTGGAGGCGCGGAGGCTCTCAAAACAGGGGAAGTTCACAGCAAGTGGATGTC 1478
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1276 ATCAAGTGGACAGCCCGAGGAGCCCTCTATGGCCGTTTACCATCAAGTGGAGTGC 1335
OY 1479 TGGAGTTTGGGGTCTGCTGTGGAGGTCTTCTATATGACGGGCTCCGTACCTTAA 1538
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1336 TGGTCTTGGCATCTGCTGTAGCTGAGTGAACCAAGGGCGGGTGCATACCCAGGG 1395
OY 1539 ATGTCACTGAAGAGGTGCGAGGCGCTGAGAAAGGGGTACCGCATGGAACCCCGAG 1598
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1396 ATGCTCAACAGGAGGTGCTGAGCAGGTGAGAGGGGCTACCGCATGCGCTGCGCGCC 1455
OY 1599 GGGTGTCCAGGCGCGGAGTCTCATGAGCAGCTGTGGAGGAGAGCCCGCCGCGC 1658
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1456 GAGTCCCGGAGTCCCTGATATGACTCATGTGCTGCGGAGGAGCCCTTGAAGAG 1515
OY 1659 CGGCCACCTTCCGCAAACTGGCCGAGAAAGCTGG 1692
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1516 CGGCCACCTTTTGAATACCTGCAAGGCTTCTGG 1549
```

Search completed: August 1, 2003, 19:44:23
Job time : 187 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 13:32:53 ; Search time 669 Seconds
(without alignments)
8070.069 Million cell updates/sec

Title: US-09-977-260-1

Perfect score: 2000

Sequence: 1 ctgcctcaagtgtgcagc.....attctaagactctaaataa 2000

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03:.*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1998.4	99.9	2000	AAT00616	Megakaryocyte kina
2	1963.4	98.2	1997	AAL50631	Human matk gene-re
3	1909.8	95.5	1987	AAV44497	Human matk CDNA.
4	1901.8	95.1	1942	AAO84888	DNA encoding cytop
5	1377	68.8	1713	ABK88791	cDNA encoding huma
6	547.6	27.4	2187	AAZ28701	Wild-type human c-
7	547.6	27.4	2187	AAH28358	Nucleotide sequenc
8	547.6	27.4	2187	ABK84602	Human CDNA differe

9	547.6	27.4	2187	24	ABL62918
10	547.6	27.4	2420	21	AAZ46489
11	547.6	27.4	2420	25	ACC50120
12	547.6	27.4	2466	22	AAI93804
13	527.8	26.4	2442	24	ABK73326
14	455	22.8	16389	24	ABK88792
15	330.2	16.5	2591	23	ABK11799
16	213.6	10.7	3299	22	AAH18556
17	213.6	10.7	4466	24	ABN59752
18	212.4	10.6	1090	23	AAK87964
19	212.4	10.6	1611	14	AAQ46688
20	212.4	10.6	1611	24	ABK73325
21	212.4	10.6	1611	25	ABZ59382
22	212.4	10.6	1699	23	AAK87965
23	204.2	10.2	2282	24	ABK56203
24	202.6	10.1	2032	21	AAZ4491
25	199.2	10.0	1602	14	AAQ46687
26	197.6	9.9	1759	21	AAZ29700
27	197.6	9.9	1759	22	AAH28357
28	193.4	9.7	1626	25	ABZ59389
29	192.6	9.6	1926	24	ABK83940
30	192.6	9.6	2015	24	ABK83939
31	192.6	9.6	2015	24	ABL66673
32	191.4	9.6	3323	25	ABZ18519
33	191.4	9.6	3380	25	ABZ18520
34	191.4	9.6	5520	18	AAK61865
35	191.4	9.6	5527	24	ABN97215
36	186.2	9.3	1911	24	ABK63704
37	185	9.2	1574	21	AAZ86794
38	185	9.2	1574	22	AAZ1845
39	185	9.2	1574	24	ABO76288
40	185	9.2	3527	23	ABK66986
41	185	9.2	7487	23	AAK92457
42	184.6	9.2	2354	24	ABK83948
43	184.6	9.2	2354	24	ABL68108
44	184.6	9.2	2433	24	AAK94859
45	184.4	9.2	1254	12	AAO13983

ALIGNMENTS

RESULT 1	AAAT00616	standard: CDNA; 2000 BP.
AC	AAAT00616	
AC	AAAT00616	
DT	26-MAR-1996	(first entry)
XX		
DE	Megakaryocyte kinase MKK1 CDNA.	
XX		
KW	Megakaryocyte kinase-1; MKK1; cytoplasmic tyrosine kinase;	
KW	cellular signal transduction; leukaemia; myelosis; myelofibrosis;	
XX	gene therapy; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	258..1781
FT		/*tag= a
XX		
PN	W09529185-A1.	
XX		
PD	02-NOV-1995.	
XX		
PF	24-APR-1995;	95WO-US05008.
XX		
PR	21-APR-1995;	95US-0426509.
XX		
PR	22-APR-1994;	94US-0232545.
XX		
PA	(PLAC) MAX PLANC GES FOERDERUNG WISSENSCHAFTEN.	
XX		
PA	(SUGE-) SUGEN INC.	

Breast cancer rela
PKA substrate, Csk
Breast cancer asso
Human polynucleoti
DNA encoding human
Gene encoding human
Drosophila melanog
Human CDNA sequenc
Novel human coding
DNA encoding novel
Human p60 c-src 9
DNA encoding human
Human src-c encodi
DNA encoding novel
cDNA encoding tumo
PKA substrate, Src
Chicken p60 c-src
Wild-type chicken
Nucleotide sequenc
Mouse src-c encodi
Human CDNA differe
Human CDNA differe
Lung cancer relate
Group III CDNA can
c-abl gene. Homo
Gene #3713 used to
Rat sequence diffe
Human protein kina
Human protein kina
Human protein kina
DNA encoding novel
DNA encoding novel
Human CDNA differe
Ovary cancer relat
Human DNA sequence
Lck gene fused wit

XX Gishizky M, Sures I, Ullrich A:
XX WPI: 1995-382959/49.
DR P-PSDB: AAR84181.
XX New poly:nucleotide(s) encoding megakaryocyte tyrosine kinase(s) -
PT used to develop prods. for the treatment and diagnosis of kinase
PT related signal transduction abnormalities.
XX
XX Claim 2: Fig 1A-C; 82pp; English.
XX
XX Overlapping cDNA clones were combined to obtain a sequence (AAT00616)
CC coding for human megakaryocyte kinase MKK1 (AAR84181). The clones
CC were isolated from a fetal brain library using degenerate primers
CC (AAT00614-15) based on conserved regions within the kinase domain of
CC receptor tyrosine kinases. MKK polynucleotides can be used in the
CC procdn. in host cells of recombinant MKK, and in the gene therapy of
CC diseases such as acute megakaryocytic leukemia, myelofibrosis and
CC acute megakaryocytic myelosis.
XX
SQ Sequence 2000 BP; 375 A; 631 C; 674 G; 320 T; 0 other;
Query Match 99.9%; Score 1998.4; DB 16; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCGCTCCAAAGTTTGTCAGCCGCGGAGCCCTCGGGGTGTCAGCCGCGTGGAGGCCC 60
DB 1 CTCGCTCCAAAGTTTGTCAGCCGCGGAGCCCTCGGGGTGTCAGCCGCGTGGAGGCCC 60
QY 61 TCCGCGGG 120
DB 61 TCCGCGGG 120
QY 121 AGGCTCGGTCCAGTGGGACCCAGCTCCCTACCTCTCTGTCAGAGCGGCTGAGCCGTGGA 180
DB 121 AGGCTCGGTCCAGTGGGACCCAGCTCCCTACCTCTCTGTCAGAGCGGCTGAGCCGTGGA 180
QY 181 GGGCATTTCCCAAGGTCGCCACGTGTACCACTTGTCTCAGTGTGCTCTCACTGCTCAG 240
DB 181 GGGCATTTCCCAAGGTCGCCACGTGTACCACTTGTCTCAGTGTGCTCTCACTGCTCAG 240
QY 241 TTTCTCTCTGCGGGGGAGTGG 300
DB 241 TTTCTCTCTGCGGGGGAGTGG 300
QY 301 GCTGTGATTTCTGCTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCAGCCGTGGGACC 360
DB 301 GCTGTGATTTCTGCTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCAGCCGTGGGACC 360
QY 361 CCCCTCCCGTCTAGCGAAGATGCCAAGAGGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGG 420
DB 361 CCCCTCCCGTCTAGCGAAGATGCCAAGAGGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGG 420
QY 421 CCAAATGTCAGACACCCCGCCCAAGCCAGGAGGAGTGGGCTTCCGAAAGGGGAGAGTGG 480
DB 421 CCAAATGTCAGACACCCCGCCCAAGCCAGGAGGAGTGGGCTTCCGAAAGGGGAGAGTGG 480
QY 481 TCACATCTCTGAGAGGCTGCGAGAACAGAGTGTACCGCTCAAGCACACACAGTGG 540
DB 481 TCACATCTCTGAGAGGCTGCGAGAACAGAGTGTACCGCTCAAGCACACACAGTGG 540
QY 541 GACATGAGGGGGGCTGCGAGAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 600
DB 541 GACATGAGGGGGGCTGCGAGAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 600
QY 601 CCNAAGCTCAAGCTCATGCGTGGTTCACAGGGAAGATCTCGGGCCAGAGAGGCTGTCCAG 660
DB 601 CCNAAGCTCAAGCTCATGCGTGGTTCACAGGGAAGATCTCGGGCCAGAGAGGCTGTCCAG 660
QY 661 AGCTGACGCTCCCGAGATGGGCTGTCTCTGCTGGGAGTCCCGCGCCACCCCGGCG 720
DB 661 AGCTGACGCTCCCGAGATGGGCTGTCTCTGCTGGGAGTCCCGCGCCACCCCGGCG 720

DB 661 AGCTGACGCTCCCGAGATGGGCTGTCTCTGCTGGGAGTCCCGCGCCACCCCGGCG 720
QY 721 ACTAGTCTCTGCTGAGTGGTGG 780
DB 721 ACTAGTCTCTGCTGAGTGGTGG 780
QY 781 ACGGCCACCTTCACATGATGAGGCGGCTTCTTCTGCAACCTCATGACATGTGTGAGC 840
DB 781 ACGGCCACCTTCACATGATGAGGCGGCTTCTTCTGCAACCTCATGACATGTGTGAGC 840
QY 841 ATTACAGCAAGAGGAGGCGGCTATCTGCAACCAAGCTGTGTAGACCAAGAGGAAACAG 900
DB 841 ATTACAGCAAGAGGAGGCGGCTATCTGCAACCAAGCTGTGTAGACCAAGAGGAAACAG 900
QY 901 GGACCAAGTGGCGGAGAGAGGAGCTGGCCACAGGCGGCGGCTGTCTGTAACCTGTGAGATT 960
DB 901 GGACCAAGTGGCGGAGAGAGGAGCTGGCCACAGGCGGCGGCTGTCTGTAACCTGTGAGATT 960
QY 961 TGACATTTGGAGCACAGATCGGAGAGGAGAGATTGGAGCTGTCTGCAAGGTGAGTACC 1020
DB 961 TGACATTTGGAGCACAGATCGGAGAGGAGAGATTGGAGCTGTCTGCAAGGTGAGTACC 1020
QY 1021 TGGGGCAAAAGGTGGCGGCTGAGAAATATCAAGTGTGATGTGACAGCCCAAGGCTTCTG 1080
DB 1021 TGGGGCAAAAGGTGGCGGCTGAGAAATATCAAGTGTGATGTGACAGCCCAAGGCTTCTG 1080
QY 1081 ACGAGAGCGGCTTCATGACGAAGATGCAACAGAGAACTGTGTCCTCTGCGGCGTGA 1140
DB 1081 ACGAGAGCGGCTTCATGACGAAGATGCAACAGAGAACTGTGTCCTCTGCGGCGTGA 1140
QY 1141 TCCGTCACACAGGGGCTTACATTTGTCTGAGACAGTGTGACAGAGGCAAGGCTGTGAACT 1200
DB 1141 TCCGTCACACAGGGGCTTACATTTGTCTGAGACAGTGTGACAGAGGCAAGGCTGTGAACT 1200
QY 1201 TTTGCGGAGCCGGGGGTCGAGCCCTGTGAAACCCGCTGACGCTCTGCACTTTCTCTG 1260
DB 1201 TTTGCGGAGCCGGGGGTCGAGCCCTGTGAAACCCGCTGACGCTCTGCACTTTCTCTG 1260
QY 1261 ACGTGGCGGAGGCGATGAGTACCTGTGAGAGCAAGAGCTTGTGACACCGGACCTGGCCG 1320
DB 1261 ACGTGGCGGAGGCGATGAGTACCTGTGAGAGCAAGAGCTTGTGACACCGGACCTGGCCG 1320
QY 1321 CCCGCAACATCTGTGTTCTAGAGAGCTGTGTGGCCAAAGTGTGACGACTTTGGCCTGGCCA 1380
DB 1321 CCCGCAACATCTGTGTTCTAGAGAGCTGTGTGGCCAAAGTGTGACGACTTTGGCCTGGCCA 1380
QY 1381 AAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGCTGCCCTGCAAGTGGAGCGGCGGAGG 1440
DB 1381 AAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGCTGCCCTGCAAGTGGAGCGGCGGAGG 1440
QY 1441 CTCTCAAAACAGGGGAAGTTTACCAAGCAAGTGTGAGTGTGGGGTCTCTCTCT 1500
DB 1441 CTCTCAAAACAGGGGAAGTTTACCAAGCAAGTGTGAGTGTGGGGTCTCTCTCTCT 1500
QY 1501 GGGAGGCTTCTCATATGAGAGGGGCTCCGTAACCTTAAATGTACGTAAGAGAGTGTGG 1560
DB 1501 GGGAGGCTTCTCATATGAGAGGGGCTCCGTAACCTTAAATGTACGTAAGAGAGTGTGG 1560
QY 1561 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGCTGTCCAGGCGCGCTGACAG 1620
DB 1561 AGGCGGTGAGAGAGGGGTACCGCATGGAACCCCGGAGGCTGTCCAGGCGCGCTGACAG 1620
QY 1621 TCTCTATGAGACGCTCTGTGGAGGACAGAGCGGCGCGGCGGCGGCGGCGGCGGCGG 1680
DB 1621 TCTCTATGAGACGCTCTGTGGAGGACAGAGCGGCGCGGCGGCGGCGGCGGCGGCGG 1680
QY 1681 CCGAGAAGCTGGCGGCGGAGCTAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1740
DB 1681 CCGAGAAGCTGGCGGCGGAGCTAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1740
QY 1741 ACGCGAGCGGCTTCACACTGCGCCCGGAGAGGAGGCTGTGACCCGACCCGCGGTGGGCGCT 1800
DB 1741 ACGCGAGCGGCTTCACACTGCGCCCGGAGAGGAGGCTGTGACCCGACCCGCGGTGGGCGCT 1800

QY	1801	TGGCCCCAGAGACCGAGAGGTGAGAGTGTGGCGTGGGGGGCACTAGCCAGGCCCAAGG	1860
Db	1801	TGGCTCCAGAGACCGAGAGGTGAGAGTGTGGCGTGGGGGGCACTAGCCAGGCCCAAGG	1860
QY	1861	AGGGTCCAGCGGGGCAAGTCATCCTCGTGGTGCCACACAGCAGGGGGCTGGGCCACGTAGGG	1920
Db	1861	AGGGTCCAGCGGGGCAAGTCATCCTCGTGGTGCCACACAGCAGGGGGCTGGGCCACGTAGGG	1920
QY	1921	GGCTCTGGGGGCCCCGTGGACACCCCAAGCTGCGAAGATGATGCCCGATAAAGCGG	1980
Db	1921	GGCTCTGGGGGCCCCGTGGACACCCCAAGCTGCGAAGATGATGCCCGATAAAGCGG	1980
QY	1981	ATTCTAAGGACTCTAAAAA 2000	
Db	1981	ATTCTAAGGACTCTAAAAA 2000	

```

RESULT 2
AAL50631
ID AAL50631 standard; DNA; 1997 BP

```

DT 02-JAN-2003 (first entry)

Human matk gene-related coding sequence.

KW Human; gene; ds; allergic disease; matk gene; atopic dermatitis

OS Homo sapiens.

FH	key	Location/Qualifiers
FH	key	Location/Qualifiers

```
/*tag= a
```

XX

XX
117

2000

XX
FE
01-MAR-2002; 2002M0-0F01910.

XX
XX
XX

0701900-3CT00Z
T00Z - WMM - T2

PA	(GENU-)	GENOX	RES	INC.
PA	(NIGE-)	JAPAN	GEN	AGENCY NATION

xx Sugita Y, Helshl M, Kagaaya S, Gunji S, Tsujimoto G;
PI

XX
DR WPI: 2002-750572/81.

DR P-PSDB; AAOL3600.
XX

PT Examining allergic diseases by changes in expression level of mark gene
PT in peripheral monocytes as indication, also applicable in screening

PI compounds for treating allergic disease

ps : Disclosure; Page 48-52; 60pp; Japanese.
XX

CC The invention comprises a method of examining allergic diseases, the
CC method comprises comparing the expression level of the mark gene in a

biological sample of a patient and a healthy individual. The method of the invention is useful for examining/diagnosing allergic diseases -

CC particularly atopic dermatitis, the method of the invention is also
CC useful in screening candidate compounds for remedies. The present DNA

cc sequence encodes a human matk-related protein.

Sequence 1997 BP; 372 A; 631 C; 676 G; 318 T; 0 other; 50

Query Match	98.28;	Score 1963.4;	DB 24;	Length 1997;
Best Local Similarity	99.88;	Pred NO:		

Matches 1987; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY	1	TTGCGTCCAAATTGGTGCAGCCGGGAAACCGCCCTGGGGGTGTGCAGCGCCCTGGCGGAGGCC	60
Db	8	CTGCTGCCAAATTGGTGCAGCCGGGAAACCGCCCTGGGGGTGTGTGAGCGGCTGCGGAGGCC	67
QY	61	TCCTGGGGGGCGGGGGCGGG--GGGGCTTCGGGGGGGCCCTTGAGCAGAAAAACAGAGAA	118
Db	68	TCCTGGGGGGCGGGGGCGGGGGGGGGCTCTGGGGGGGCCCTTGAGCAGAAAAACAGAGAA	127
QY	119	CCAGGCTTCGGTCCAGTGGACCCAGCTCCCTACCTCTGTGCGACGCGGCTGGCTGTGG	178
Db	128	CCAGGCTTCGGTCCAGTGGACCCAGCTCCCTACCTCTGTGCGACGCGGCTGGCTGTGG	187
QY	179	CAGGCGCAATCCCAAGCTCCCGGACTGTGACCACTTGTCTCAATGTGGCTCTACCTGGCTC	238
Db	188	CAGGCGCAATCCCAAGCTCCCGGACTGTGACCACTTGTCTCAATGTGGCTCTACCTGGCTC	247
QY	239	AGTTTCCCTTCGGGGGGGCGATGGGCGAGGCTCTGTGGTTTCTTGCGCGGCAATTTC	298
Db	248	AGTTTCCCTTC--GGGGGGGATGGGGGGGCGAGGCTCTGTGGTTTCTTGCGCGGCAATTTC	306
QY	299	CGGCTGTGATTTCTGCTGAGAACATTCCCCGGGTGAGCCCCCGCTTCTCCGAGCCTGGCA	358
Db	307	CGGCTGTGATTTCTGCTGAGAACATTCCCCGGGTGAGCCCCCGCTTCTCCGAGCCTGGCA	366
QY	359	CCCCCTCCCTCTCAGCCAGAGATGCCAACAGAGCGCTGGGGCCCCCGGCAACCCAGTGTAT	418
Db	367	CCCCCTCCCTCTCTCAGCCAGAGATGCCAACAGAGCGCTGGGGCCCCCGGCAACCCAGTGTAT	426
QY	419	CACCAAAATGCAGAGCACACCCGCCCAAGCCAGGGAGCTGTGCTTCGCAAGGGCGAGT	478
Db	427	CACCAAAATGCAGAGCACACCCGCCCAAGCCAGGGAGCTGTGCTTCGCAAGGGCGAGT	486
QY	479	GGTCACCATCTCTGAGGCGCTGCGAAGAACAGAGCTGGTACCGGCTCAAGCACACACACAG	538
Db	487	GGTCACCATCTCTGAGGCGCTGCGAAGAACAGAGCTGGTACCGGCTCAAGCACACACACAG	546
QY	539	TGACAGAGAGGGGCGCTGTGCGACGCTGGGGGGCTGCGGAGCGGGAGGCCCTCTCCGAGA	598
Db	547	TGAGAGAGAGGGGCGCTGTGCGACGCTGGGGGGCTGCGGAGCGGGAGGCCCTCTCCGAGA	606
QY	599	CCCCAAGTCAAGCCTCATGCCGTGGTTCCAGGGGAAGTCTCGGGGCCAGAGGCGTGTCCA	658
Db	607	CCCCAAGTCAAGCCTCATGCCGTGGTTCCAGGGGAAGTCTCGGGGCCAGAGGCGTGTCCA	666
QY	659	GCAGCTGCAGCCTCCCGAGAGTGGGCTTTCCTGTGGGGAGTCCGCGGCCACCCCGG	718
Db	667	GCAGCTGCAGCCTCCCGAGAGTGGGCTTTCCTGTGGGGAGTCCGCGGCCACCCCGG	726
QY	719	CGACTAGCTCTGTGCGTGTGAGCTTTGGCGCGAGCTCATCCACTACCGCGCTGTGCACCG	778
Db	727	CGACTAGCTCTGTGCGTGTGAGCTTTGGCGCGAGCTCATCCACTACCGCGCTGTGCACCG	786
QY	779	CGAGCGGCCACCTCACAATCGATGAGGCCGTGTCTTCGCAACCTCATGTGCATGTGTGA	838
Db	787	CGAGCGGCCACCTCACAATCGATGAGGCCGTGTCTTCGCAACCTCATGTGCATGTGTGA	846
QY	839	GCATTACGACAGCAAGAGGGCGGTATCTGCACCAAGCTGTGTGAGACCCAAAGCGGAACA	898
Db	847	GCATTACGACAGCAAGAGGGCGGTATCTGCACCAAGCTGTGTGAGACCCAAAGCGGAACA	906
QY	899	CGGCAACAAGTGGCGCCAGAGAGAGCTGTGCGAGGGCGGCTGTACTGAACCTGTGCACA	958
Db	907	CGGCAACAAGTGGCGCCAGAGAGAGCTGTGCGAGGGCGGCTGTACTGAACCTGTGCACA	966
QY	959	TTTGCATTTGGGAGACAGATCGGAGAGAGAGTTTGGACTGTGCTTCGACGGGTGAGTA	1018
Db	967	TTTGCATTTGGGAGACAGATCGGAGAGAGAGTTTGGACTGTGCTTCGACGGGTGAGTA	1026
QY	1019	CCTGGGGCAAAAGGTGGCCGTGAAGAATATCAAGTGTGATGTGACAGCCAGGCTTCTCT	1078
Db	1027	CCTGGGGCAAAAGGTGGCCGTGAAGAATATCAAGTGTGATGTGACAGCCAGGCTTCTCT	1086
QY	1079	GGAGGAGAGCGCCGTATGACGAAGATGCAACACGACACCTGGTGGCTTCTCTGGGCGT	1138

```

Db 1087 GGAGAGAGAGGCGCTCATGACGAAGATGCAACAGAGAACTGGTGGCTCTCTGGGCGT 1146
QY 1139 GATCTTCACACAGGGGGGTGTATCATTTTCATGAGACAGTGAAGGCAACCTGTGTAA 1198
Db 1147 GATCTTCACACAGGGGGGTGTATCATTTTCATGAGACAGTGAAGGCAACCTGTGTAA 1206
QY 1199 CTTTCTGCGAGACCGGGGGGTGAGCGCTGCTGTAACACCGCTGAGTCTGCACTTTTCTCT 1258
Db 1207 CTTTCTGCGAGACCGGGGGGTGAGCGCTGCTGTAACACCGCTGAGTCTGCACTTTTCTCT 1266
QY 1259 GCAGTGGCGAGAGGAGTGTAGTACCTGAGAGAGCAAGAGTTGTGCACCGCGACCTGGC 1318
Db 1267 GCAGTGGCGAGAGGAGTGTAGTACCTGAGAGAGCAAGAGTTGTGCACCGCGACCTGGC 1326
QY 1319 CGCCCGCAACATCTGTCTCAGAGAGACCTGTGGGCCAAGTCAAGCACTTTTGGCTGGC 1378
Db 1327 CGCCCGCAACATCTGTCTCAGAGAGACCTGTGGGCCAAGTCAAGCACTTTTGGCTGGC 1386
QY 1379 CAAAGCGGAGGAGGAGGCTAGACCTCAAGCGGCTCCGCTCAAGTGGAGCGGCGCGA 1438
Db 1387 CAAAGCGGAGGAGGAGGCTAGACCTCAAGCGGCTCCGCTCAAGTGGAGCGGCGCGA 1446
QY 1439 GCGCTCTCAACACAGGGAAGTTCAACCAAGTGGATGTCTGAGAGTTTGGGGTGTCT 1498
Db 1447 GCGCTCTCAACACAGGGAAGTTCAACCAAGTGGATGTCTGAGAGTTTGGGGTGTCT 1506
QY 1499 CTGGGAGGTTTCTCATATGAGAGCGGCTCGTACCTTAAATGTCACTGAAAGAGTGTCT 1558
Db 1507 CTGGGAGGTTTCTCATATGAGAGCGGCTCGTACCTTAAATGTCACTGAAAGAGTGTCT 1566
QY 1559 GGAGGCGGTGAGAGAGGAGTACCGCAAGCAAGCAAGCGGAGGCTGTCCAGCGCGCGTGA 1618
Db 1567 GGAGGCGGTGAGAGAGGAGTACCGCAAGCAAGCAAGCGGAGGCTGTCCAGCGCGCGTGA 1626
QY 1619 CGTCTCTATGAGAGTGTCTGAGAGAGCGGCGCGCGCGCGCACTTCCGCAACT 1678
Db 1627 CGTCTCTATGAGAGTGTCTGAGAGAGCGGCGCGCGCGCGCGCACTTCCGCAACT 1686
QY 1679 GGCGGAGAACTGGCGCGGAGAGTACGCAAGTGTGCGCGCAAGCTCGCTTCAGGGCA 1738
Db 1687 GGCGGAGAACTGGCGCGGAGAGTACGCAAGTGTGCGCGCAAGCTCGCTTCAGGGCA 1746
QY 1739 GGAGGCGGAGGCTCTCACTCGCGCGCGCGCGCGCGCGCGCACTTCCGCAACT 1798
Db 1747 GGAGGCGGAGGCTCTCACTCGCGCGCGCGCGCGCGCGCGCGCACTTCCGCAACT 1806
QY 1799 CTGGGCGGAGAGAGAGAGTGTGAGAGTGTGCGGCTGGGGCACTGACCAAGGCCAA 1858
Db 1807 CTGGGCGGAGAGAGAGAGTGTGAGAGTGTGCGGCTGGGGCACTGACCAAGGCCAA 1866
QY 1859 GGAGGCGGAGGCGGCGGCAAGTCTCTCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCA 1918
Db 1867 GGAGGCGGAGGCGGCGGCAAGTCTCTCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCA 1926
QY 1919 GGCGGCTGTGGGCGCGCGGAGAGAGAGTGTGAGAGAGTGTGCGGCGCGCGCGCGCGCA 1978
Db 1927 GGCGGCTGTGGGCGCGCGGAGAGAGAGTGTGAGAGAGTGTGCGGCGCGCGCGCGCGCA 1986
QY 1979 GGATTTCTAAGG 1989
Db 1987 GGATTTCTAAGG 1997

```

RESULT 3
AAV44497
ID AAV44497 standard; cDNA; 1987 BP.
XX
AC AAV44497;
XX
DT 16-OCT-1998 (first entry)
XX
DE Human matk cDNA.

```

XX XX Breast; cancer; matk; CSK homologous kinase; CHK; detection; diagnosis;
KW KW cytoplasmic protein; tyrosine kinase; ErbB-2; negative regulator;
KW KW mitogenic signalling; ss.
OS OS Homo sapiens.
XX XX
FH FH Key Location/Qualifiers
FT FT CDS 263..1846
FT FT /*tag= a
FT FT /product= CHK
FT FT /note= "CSK homologous kinase"
PE PE W09830704-A1.
PD PD 16-JUL-1998.
PD PD
PE PE 07-JAN-1998; 98MO-US00420.
PR PR 16-JUN-1997; 97US-0876882.
PR PR 08-JAN-1997; 97US-0035228.
PA PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PI PI Avraham H, Groopman JE;
XX XX WPI: 1998-399149/34.
XX XX P-PSDB: AAM64454.
PT PT Detecting breast cancer by detecting Csk homologous kinase
PT PT expression - especially in humans and use of Csk homologous kinase
PT PT in treatment or prophylaxis of breast cancer and for producing
XX XX medicaments
XX XX
PS PS Disclosure; Fig 2; 54pp; English.
XX XX
CC CC This sequence encodes a CSK homologous kinase (CHK) which is used in a
CC CC method of detecting cancer in breast tissue. The method allows diagnosis
CC CC of breast cancer in mammals, especially humans. It is based on the
CC CC discovery that a cytoplasmic protein tyrosine kinase, CHK, is expressed
CC CC in human breast tissue, but not in adjacent tissue. This protein can be
CC CC used to raise antibodies which can be included in compositions and
CC CC diagnostic kits for diagnosis of breast cancer. The presence of CHK in
CC CC breast tissue can also be determined using other standard methods (e.g.
CC CC Northern blotting) or by detecting nucleic acid sequences encoding all/a
CC CC portion of the protein (e.g. using hybridisation probes). Over-expression
CC CC of the receptor tyrosine kinase ErbB-2 has previously been associated
CC CC with the development of breast cancer, and CHK specifically interacts
CC CC with activated ErbB-2, and may function as a negative regulator of
CC CC ErbB-2 mediated mitogenic signalling. The compositions may also be used
CC CC to design drugs (e.g. which incorporate CHK analogues with greater
CC CC biological activity than CHK) and to identify CHK antagonists and
CC CC agonists for therapeutic use.
XX XX
SO SO Sequence 1987 BP; 369 A; 628 C; 672 G; 318 T; 0 other;
QY QY
Query Match 95.5%; Score 1909.8; DB 19; Length 1987;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1973; Conservative 0; Mismatches 7; Indels 9; Gaps 5;
Db 1 CTGCGTCCAAAGTGTGACAGCGGAGCGCTGCGGGGTGTGACAGCGGCTGCGGAGGCC 60
Db 8 CTGCGTCCAAAGTGTGACAGCGGAGCGCTGCGGGGTGTGACAGCGGCTGCGGAGGCC 67
QY 61 TCCTGGGGGCGGCGGCGGCGGCTGCGGGGCGCGCGCTGAGAGCAAAACAGGAACCC 120
Db 68 TCCTGGGGGCGGCGGCGGCGGCGCTGCGGGGCGCGCGCTGAGAGCAAAACAGGAACCC 127
QY 121 AGGCTGGGTCCAGTGTGACAGCGCTTCTGCTGTGCAAGCGCGCTGTGGCA 180
Db 128 AGGCTGGGTCCAGTGTGACAGCGCTTCTGCTGTGCAAGCGCGCTGTGGCA 187
QY 181 GGCAATTCACAGCGTCCCGAGCTGTGACCACTTGTCTGAGTGTGCTTCACTGCTCTAG 240

```

188 GGCATTCCACGCGCCGACCTGTGACACTTGTGCTCAGTGTGCTTACCTGCTGCTAG 247
 241 TTTCCTCTGGGGGGGCGATGGGGGGGCGAGGCTCTGTGTTCCTGGGGGGGATTTACG 300
 248 TTTCC--TCGTGGGGGGATGGGGGGGCGAGGCTCTGTGTTCCTGGGGGGGATTTACG 305
 301 GCTGTGATTTGCTGTAGGAACCTTCCCGGGTGAGCCCCGGCTTCTTCCGAGCCTGGAC 360
 306 GCTGTGATTTGCTGTAGGAACCTTCCCGGGTGAGCCCCGGCTTCTTCCGAGCCTGGAC 365
 361 CCCCTCCGCTCTACGCCAGAGATGCCACGAGGCGCTGGGCCGCCGAGGACCCAGTGTATCA 420
 366 CCCCTCCGCTCTACGCCAGAGATGCCACGAGGCGCTGGGCCGCCGAGGACCCAGTGTATCA 425
 421 CCAATGTGAGACACCGCCCGCCAGCCGAGGAGCTGGCTTCCGCAAGGGCGACGTGG 480
 426 CCAATGTGAGACACCGCCCGCCAGCCGAGGAGCTGGCTTCCGCAAGGGCGACGTGG 485
 481 TCACATCTGTGAGGCTGTGCGAACAAGAGCTGTACCGGCTCAAGCACACACAGTGTG 540
 486 TCACATCTGTGAGGCTGTGCGAACAAGAGCTGTACCGGCTCAAGCACACACAGTGTG 545
 541 GACAGAGAGGCGCTGTGCGACAGTGGGGCGCTGTGCGGAGCGGGAGGCGCTTTCGCGAGAC 600
 546 GACAGAGAGGCGCTGTGCGACAGTGGGGCGCTGTGCGGAGCGGGAGGCGCTTTCGCGAGAC 605
 601 CCAAGCTAGGCTCTATGCGGTGTCCCAAGGGAATATCTGGGCGCCAGAGAGGTGTCCAGC 660
 606 CCAAGCTAGGCTCTATGCGGTGTCCCAAGGGAATATCTGGGCGCCAGAGAGGTGTCCAGC 665
 661 AGCTGACAGCTTCCCGAGAGATGGGCTGTCTGTGTCGCGGAGTCCGCGGCCACCCGCGCG 720
 666 AGCTGACAGCTTCCCGAGAGATGGGCTGTCTGTGTCGCGGAGTCCGCGGCCACCCGCGCG 725
 721 ACTAGCTCTGTGCTGAGCTTTTGGCCGCGAGCTATCCACTACCGCGCTGTGACCGCG 780
 726 ACTAGCTCTGTGCTGAGCTTTTGGCCGCGAGCTATCCACTACCGCGCTGTGACCGCG 785
 781 ACGGCGACCTCAACATGATGAGGCGGCTTGTGTAACCTCATGTGACATGTGTGAGC 840
 786 ACGGCGACCTCAACATGATGAGGCGGCTTGTGTAACCTCATGTGACATGTGTGAGC 845
 841 ATTACAGCAAGACAAAGGCGCTATCTGACCAAGCTGTGTGAGACCAAGGCGGAACAGC 900
 846 ATTACAGCAAGACAAAGGCGCTATCTGACCAAGCTGTGTGAGACCAAGGCGGAACAGC 905
 901 GGACCAAGTCCGCGAGAGAGAGCTGGCCAGGGCGGCTGTGTACTGAACCTGCAGCAATT 960
 906 GGACCAAGTCCGCGAGAGAGAGCTGGCCAGGGCGGCTGTGTACTGAACCTGCAGCAATT 965
 961 TGACATTGGGAGACAGATGCGAGAGGAGATTGGAGCTGTCTGTGAGGTTGATGATC 1020
 966 TGACATTGGGAGACAGATGCGAGAGGAGATTGGAGCTGTCTGTGAGGTTGATGATC 1025
 1021 TGGGGCAAAAGGTGGCCGTGAAGAATATCAAGTGTGATGTACAGGCCAGGCTTCTCTG 1080
 1026 TGGGGCAAAAGGTGGCCGTGAAGAATATCAAGTGTGATGTACAGGCCAGGCTTCTCTG 1085
 1081 ACGAGAGCGCGCTCATGTACGAAGATGCAACACAGAGACCTGTGTCTCTGGGCGTGA 1140
 1086 ACGAGAGCGCGCTCATGTACGAAGATGCAACACAGAGACCTGTGTCTCTGGGCGTGA 1145
 1141 TCTGACACGAGGCGCTGTACATTTGTCATGTGAGACAGCTGTGAGAGGCAACCTGTGATC 1200
 1146 TCTGACACGAGGCGCTGTACATTTGTCATGTGAGACAGCTGTGAGAGGCAACCTGTGATC 1205
 1201 TTTGCGGAGCCCGGGGCTGAGCCCTCTGTGAACACCGCTCAGCTCTGCAAGTTTCTCTGC 1260
 1206 TTTGCGGAGCCCGGGGCTGAGCCCTCTGTGAACACCGCTCAGCTCTGCAAGTTTCTCTGC 1265
 1261 ACGTGGCGGAGGAGTGTGATCTGTGAGAGACAGAACTTGTGACCGCGCACTGGCGG 1320

1266 ACGTGGCGGAGGAGTGTGATCTGTGAGAGCAAGAACTGTGTGACCGCGCACTGGCGG 1325
 1321 CCCGCAACATCTGTGTCACAGAGACCTGTGTGCCAAGTCAAGCACTTTGGCCGTGCCA 1380
 1326 CCCGCAACATCTGTGTCACAGAGACCTGTGTGCCAAGTCAAGCACTTTGGCCGTGCCA 1385
 1381 AAGCGAGCGGAGAGGCGGTAGACTCAACCGGCTCCCGTCAAGTGTGAGAGCGGCCGAGG 1440
 1386 AAGCGAGCGGAGAGGCGGTAGACTCAACCGGCTCCCGTCAAGTGTGAGAGCGGCCGAGG 1445
 1441 CTCTCAACACGCGGAGTTCAACAGCAAGTGTGTGAGTTTGGGGTGTGCTCT 1500
 1446 CTCTCAACACGCGGAGTTCAACAGCAAGTGTGTGAGTTTGGGGTGTGCTCT 1502
 1501 GGGAGGCTTCTCATATGTGAGCGGCTCCGTACCTTAAATGTCACTGAAAGAGTGTGG 1560
 1503 GGGAGGCTTCTCATATGTGAGCGGCTCCGTACCTTAAATGTCACTGAAAGAGTGTGG 1562
 1561 AGGCGGTGAGAGAGGAGTACCGCATGGAACCCCGAGAGGCTGTCCAGGCCCGTGCAGC 1620
 1563 AGGCGGTGAGAGAGGAGTACCGCATGGAACCCCGAGAGGCTGTCCAGGCCCGTGCAGC 1622
 1621 TCTCATGAGACAGTGTGTGGAGAGCAAGCCCGCGCGCACCTTTCGCAAACTGG 1680
 1623 TCTCATGAGACAGTGTGTGGAGAGCAAGCCCGCGCGCACCTTTCGCAAACTGG 1681
 1681 CCGAAGACCTGGCCCGGAGAGTACCGCATGGAAGTGTGCCCAACCTTCTGTAGGGCAGG 1740
 1682 CCGAAGACCTGGCCCGGAGAGTACCGCATGGAAGTGTGCCCAACCTTCTGTAGGGCAGG 1741
 1741 ACGCGGAGGCTCTACCTTCGCGCCGAGAGCAAGAGCCCTGACCCACCGCGGCGGAGCT 1800
 1742 ACGCGGAGGCTCTACCTTCGCGCGAGAGCAAGAGCCCTGACCCACCGCGGAGCT 1798
 1801 TGGCCCGAGAGAGCCGAGAGAGTGTGAGTGTGCGGCGGTGTGAGACCAAGGCCCAAG 1860
 1799 TGGCCCGAGAGAGCCGAGAGAGTGTGAGTGTGCGGCGGTGTGAGACCAAGGCCCAAG 1858
 1861 AGGCTCAGAGGCGGAGCAAGTCACTGTGTGTGCGGCGGAGAGGCGGCTGGCCACTAGG 1920
 1859 AGGCTCAGAGGCGGAGCAAGTCACTGTGTGTGCGGCGGAGAGGCGGCTGGCCACTAGG 1918
 1921 GGCCTGTGGGCGGCGCGGTGAGACCCAGACCTGTGAGAGATGATGCGCGATTAAGACGG 1980
 1919 GGCCTGTGGGCGGCGCGGTGAGACCCAGACCTGTGAGAGATGATGCGCGATTAAGACGG 1978
 1981 ATTCTAAG 1989
 1979 ATTCTAAG 1987

RESULT 4
 AA084888
 ID AA084888 standard; cDNA to mRNA; 1942 BP.
 XX
 AC AA084888;
 XX
 DT 25-MAR-2003 (updated)
 DT 26-OCT-1995 (first entry)
 DE DNA encoding cytoplasmic tyrosine kinase.
 DE cytoplasmic; tyrosine kinase; blood; cell differentiation;
 KW screening; anticancer agent; ds.
 KW
 OS Homo sapiens.
 OS
 XX
 FH Key location/Qualifiers
 FT CDS 208..1731
 FT /tag= a
 FT /product= cytoplasmic-tyrosine_kinase
 FT /note= "see AAR71133"
 FT misc_feature 349..540

FT FT /tag- b
FT /note- "encodes SH3 domain (see AAR71129)"
FT misc_feature 571..795
FT /tag- c
FT /note- "encodes SH2 domain (see AAR71130)"
FT misc_feature 904..1641
FT /tag- d
FT /note- "encodes tyrosine kinase domain (see AAR71131)"
FT misc_feature 331..1728
FT /tag- e
FT /note- "encodes N-terminally truncated form of the enzyme (see AAR71132)"
XX WO9506113-A1.
XX 02-MAR-1995.
XX 25-AUG-1994.
XX 25-AUG-1993.
XX 25-AUG-1993.
XX 29-MAR-1994.
XX (ASAH) ASAH KASEI KOGYO KK.
XX Sakano S:
XX WPI: 1995-106842/14.
XX P-PSDB: AAR71129, AAR71130, AAR71131, AAR71132, AAR71133.
XX
XX Cytoplasmic tyrosine kinase and antibody recognising it - for
PT screening chemical substances for tyrosine kinase inhibitory or
PT activating activity for use as cancer therapy
XX
XX Claim 7: Page 49-50; 58pp; English.
XX
XX This DNA encodes a cytoplasmic tyrosine kinase which has enhanced
CC expression in connection with blood cell differentiation. It was
CC isolated from the human UT-7 blood cell line. The DNA sequences and
CC antibodies raised against the enzyme, are useful for screening agents
CC for inhibiting or activating activity on the tyrosine kinase, for
CC use as anticancer agents.
CC (updated on 25-MAR-2003 to correct PN field.)
CC
SQ Sequence 1942, BP: 365 A; 615 C; 651 G; 311 T; 0 other;
Query Match 95.1%; Score 1901.8; DB 16; Length 1942;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1938; Conservative 0; Mismatches 2; Indels 5; Gaps 3;
QY 52 CGAGAGCCCTCTGAGGAGGCGGCGG--GCGGCTGCGGAGGCGCCCTGAGCAGAAA 109
DB 1 CGAGAGCCCTCTGAGGAGGCGGCGGCGGCGGCGGCTGCGGAGGCGCCCTGAGCAGAAA 60
QY 110 CAGGAAGAACAGGCTGCGTGCAGTGGACCCAGCTCCCTACTCTCTGTCAGCCGCT 169
DB 61 CAGGAAGAACAGGCTGCGTGCAGTGGACCCAGCTCCCTACTCTCTGTCAGCCGCT 120
QY 170 GGCCTGTGGAGGCACTTCCAGAGCTCCCGACTGTGACACTTGTCTCACTGTGCTCTC 229
DB 121 GGCCTGTGGAGGCACTTCCAGAGCTCCCGACTGTGACACTTGTCTCACTGTGCTCTC 180
QY 230 ACCTGCTCAGTTTCCCTCTGGGGGCGATGCGGGGCGAGGCTCTGTTTCCAGGCG 289
DB 181 ACCTGCTCAGTTTCCCTCT- GGGGCGATGAGGGGCGAGGCTCTGTTTCCAGGCG 239
QY 290 GGCATTTCAGGCTGTGATTCCTGAGTGAAGTCTCCCGGAGTGAAGCCCGCTTCTCG 349
DB 240 GGCATTTCAGGCTGTGATTCCTGAGTGAAGTCTCCCGGAGTGAAGCCCGCTTCTCG 299
QY 350 AGCGTGGCAGCCCTCCCGTCTCAGCAGAGATGCAAGAGGCGCTGGGCGCCCGGCGAC 409
DB 300 AGCGTGGCAGCCCTCCCGTCTCAGCAGAGATGCAAGAGGCGCTGGGCGCCCGGCGAC 359

QY 410 CCAAGTATACCAAAATGCGAGACACCCCGCCCAAGCCAGAGGAGCTGCTTCCGCA 469
DB 360 CCAAGTATACCAAAATGCGAGACACCCCGCCCAAGCCAGAGGAGCTGCTTCCGCA 419
QY 470 GGGGAGCTGTGATCACTCTGAGAGGCTCTGAGAGACAGAGCTGTACCGCTCAAGCA 529
DB 420 GGGGAGCTGTGATCACTCTGAGAGGCTCTGAGAGACAGAGCTGTACCGCTCAAGCA 479
QY 530 CCAACAGTGTGAGAGAGAGGAGGCTGTGAGAGTGGGGGCTGCGGAGGAGGAGGCT 589
DB 480 CCAACAGTGTGAGAGAGAGGAGGCTGTGAGAGTGGGGGCTGCGGAGGAGGAGGCT 539
QY 590 CTCGAGAGACCCCAAGCTCAGCTCATGAGCTGTGCTTCCAGAGGAGATCTCGGAGCA 649
DB 540 CTCGAGAGACCCCAAGCTCAGCTCATGAGCTGTGCTTCCAGAGGAGATCTCGGAGCA 599
QY 650 GGCCTGTCCAGACAGCTGTGAGCTCCCGAGAGATGGGCTGTCTGTGCGGAGTCCGCG 709
DB 600 GGCCTGTCCAGACAGCTGTGAGCTCCCGAGAGATGGGCTGTCTGTGCGGAGTCCGCG 659
QY 710 CCAACCCGAGGAGTACGCTGCTGTGAGCTTGGCGGAGGCTGCTGCTGCTGCTGCTGCT 769
DB 660 CCAACCCGAGGAGTACGCTGCTGTGAGCTTGGCGGAGGCTGCTGCTGCTGCTGCTGCT 719
QY 770 GCTGACAGGAGGAGGCTCCTCAATGATGAGGCGCTGTCTGCTGCAACCTCATGGA 829
DB 720 GCTGACAGGAGGAGGCTCCTCAATGATGAGGCGCTGTCTGCTGCAACCTCATGGA 779
QY 830 CATGTGTGAGCATTTACAGCAGAGCAAGAGGCGCTATCTGCAACAGCTGTGAGACCAA 889
DB 780 CATGTGTGAGCATTTACAGCAGAGCAAGAGGCGCTATCTGCAACAGCTGTGAGACCAA 839
QY 890 GCGGAAACAGGAGGAGCAAGTCCGCGAGAGAGGAGTGGCCAGGCGGAGCTGTACTGAA 949
DB 840 GCGGAAACAGGAGGAGCAAGTCCGCGAGAGAGGAGTGGCCAGGCGGAGCTGTACTGAA 899
QY 950 CCGTGAAGCATTTGACATTTGGAGACACAGATCGAGAGAGGAGATTTGAGACTCTCTGCA 1009
DB 900 CCGTGAAGCATTTGACATTTGGAGACACAGATCGAGAGAGGAGATTTGAGACTCTCTGCA 959
QY 1010 GCGTGAAGCATTTGAGGAGCAAGAGTGGCGCTGGAAGATATCAAGTGTGATGACACCCA 1069
DB 960 GCGTGAAGCATTTGAGGAGCAAGAGTGGCGCTGGAAGATATCAAGTGTGATGACACCCA 1019
QY 1070 GGCCTTCTGAGAGAGAGGCGCTCATGAGAGATGCAACAGAGAACTGTGAGCTT 1129
DB 1020 GGCCTTCTGAGAGAGAGGCGCTCATGAGAGATGCAACAGAGAACTGTGAGCTT 1079
QY 1130 CCGTGGCGGTGATCTGACACAGGAGGCTGTACATTTGTATGAGACACGTGAGCAAGGCAA 1189
DB 1080 CCGTGGCGGTGATCTGACACAGGAGGCTGTACATTTGTATGAGACACGTGAGCAAGGCAA 1139
QY 1190 CCGTGGTGAATCTTCTGCGAGACCCGAGGCTGAGAGCTGTGTAACACCGCTTCTGCTGCA 1249
DB 1140 CCGTGGTGAATCTTCTGCGAGACCCGAGGCTGAGAGCTGTGTAACACCGCTTCTGCTGCA 1199
QY 1250 GTTTTCTGACAGTGGCGAGGAGGAGTGAAGTACCTGAGAGAGCAAGAGTTTGTGACCG 1309
DB 1200 GTTTTCTGACAGTGGCGAGGAGGAGTGAAGTACCTGAGAGAGCAAGAGTTTGTGACCG 1259
QY 1310 CGACCTGGCGCGCCGCAACATCTGTCTGAGAGACCTGTGAGGCAAGGTCAGGACATT 1369
DB 1260 CGACCTGGCGCGCCGCAACATCTGTCTGAGAGACCTGTGAGGCAAGGTCAGGACATT 1319
QY 1370 TGGCTGTGGCCAAAGCGAGCGAGAGGAGGCTGAGACTTAACCGCGCTCCGCTCAAGTGAAC 1429
DB 1320 TGGCTGTGGCCAAAGCGAGCGAGAGGAGGCTGAGACTTAACCGCGCTCCGCTCAAGTGAAC 1379
QY 1430 GGGCGCCGAGAGCTCTCAAAACAGCGAGAGTTCACACAGTGTGAGATGTGTGAGATTGAG 1489
DB 1380 GGGCGCCGAGAGCTCTCAAAACAGCGAGAGTTCACACAGTGTGAGATGTGTGAGATTGAG 1439

QY	1490	GTGTCGTCGTCGGGAGGCTCTCATATATGACGCGGGCTCCGATACCTTAATATGACAGAA	154.9
Db	1440	GGTCTGCTCTCGGAGAGCTTCTCTATATGACGGGCTCCGTAACCTTAATAATGTCACTGAA	149.9
QY	1550	AGAGGTGTCGGAGGCGCTGTGAGAAAGGGGTACCCGATGGAAACCCCCGAGGGCTGTCCAGG	160.9
Db	1500	AGAGGTGTGCGAGGCGCTGTGAGAAAGGGGTACCCGATGGAAACCCCCGAGGGCTGTCCAGG	155.9
QY	1610	CCCGGTGCACGTCCTCATATGACACTGCTGGGAGGACAGACGCCCGCCGCGGCCACACCTT	166.9
Db	1560	CCCGGTGCACGTCCTCATATGACACTGCTGGGAGGACAGAGGCCCGCCCGGCCACACCTT	161.9
QY	1670	CCGCAAACTGCGCCAGAGAACTGTGGCCCGGAGACTACGAGTGCAGGTGCCCCAGCCTCCGT	172.9
Db	1620	CCGCAAACTGCGCCAGAGAACTGTGGCCCGGAGACTACGAGTGCAGGTGCCCCAGCCTCCGT	167.9
QY	1730	CTCAGGGCAGAGACGCCGACGGCTCTCACTGCGCCCCGAGAACCGAGAGCCCTTACCCACACC	178.9
Db	1680	CTCAGGGCAGAGACGCCGACGGCTCTCACTGCGCCCCGAGAACCGAGAGCCCTTACCCACACC	173.9
QY	1790	GGTGGGGGCGCTTGGGCGCCCGACAGAGACCGAGAGTGTGAGAGTGCAGCGCGTGGGGGCACTGAC	184.9
Db	1740	GGT--GGGCGCTTGGGCGCCCGACAGAGACCGAGAGTGTGAGAGTGTGGGGGCACTGAC	179.9
QY	1850	CAGGCCCAAGAGAGGCTCCAGGCGGGGCAAGTATCTCTCTGCTGCCACAGCAGGAGGCTGG	190.9
Db	1798	CAGGCCCAAGAGAGGCTCCAGGCGGGGCAAGTATCTCTCTGCTGCCACAGCAGGAGGCTGG	185.7
QY	1910	CCCAAGTAGAGGGGCTGTGGGCGGGCCGAGACCCGAGACCTGCGAAGAGTGTGATTCGGCC	196.9
Db	1858	CCCAAGTAGAGGGGCTGTGGGCGGGCCGAGACCCGAGACCTGCGAAGAGTGTATCGCC	191.7
QY	1970	GATTAAGACGAGATTCTAAGCACTCT	199.4
Db	1918	GATTAAGACGAGATTCTAAGCACTCT	194.2

RESULT 5	
ABK88791	
ID	ABK88791 standard; cDNA; 1713 BP.
XX	
AC	ABK88791;
XX	
DT	21-OCT-2002 (first entry)
XX	
DE	cDNA encoding human kinase related to tyrosine kinase family
XX	
KW	Human; kinase; tyrosine kinase; kinase mediated disorder;
KW	haematopoietic disorder; cell signal transduction disorder;
KW	cancer; haemostatic; cytostatic; chromosome 19; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	223..1470
FT	/tag= a
FT	/product= "Kinase"
XX	
PN	WO200252018-A2.
XX	
PD	04-JUL-2002.
XX	
PF	19-DEC-2001; 2001WO-US48546.
XX	
PR	21-DEC-2000; 2000US-0741154.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Beasley EM, Shao W, Ketchum K, Di Francesco V;
XX	
DR	WPI; 2002-583568/62.
DR	P-PSDB; AAU99575.
XX	

PT	New isolated human kinase proteins and genes, useful in developing
PT	drugs, as well as for diagnosing, preventing or treating disorders
PT	associated with defective cell signal transduction, e.g. cancer or
PT	haematopoietic disorders
XX	
PS	
CC	Claim 4: Fig 1A; 78pp; English.
CC	
CC	The present invention relates to the isolation of a novel human kinase
CC	related to the tyrosine kinase family, and polynucleotide sequences
CC	encoding it. The gene encoding the kinase of the invention maps to
CC	chromosome 19. The kinase may be used for identifying a modulator of
CC	the kinase, an agent that binds to the kinase, or for identifying
CC	other members of the family. The kinase may also be used to raise
CC	antibodies which may be used in immunoassays or drug screening assays.
CC	The sequences of the invention may be used for treating a disease or
CC	condition mediated by a human kinase such as haematopoietic disorders
CC	and other disorders associated with defective cell signal transduction
CC	(e.g. cancer). The polynucleotide sequences encoding the kinase are
CC	useful for isolating and purifying the kinase, and as probes or primers.
CC	The present sequence encodes the human kinase of the invention.
XX	
SO	Sequence 1713 BP; 374 A; 499 C; 575 G; 265 T; 0 other:
Query Match	68.8%; Score 1377; DB 24; Length 1713;
Best Local Similarity	99.3%; Pred. No. 2.1e-266;
Matches 1383; Conservative	0; Mismatches 10; Indels 0; Gaps 0;
OY	608 CAGCCTCATGCCGCTGCTCCACAGGGAATCTCGGGCCAGAGAGCTGTCCACACTGCA 667
DB	297 CAGCAGCTTTCGTGGCTTCCACAGGGAATCTCGGGCCAGAGAGCTGTCCACACTGCA 356
OY	668 GCCTCCCGAGGATGGGCTTCTCTGGTCCGGGAGTCCGCGCCACCCCGGAGACTACGT 727
DB	357 GCCTCCCGAGGATGGGCTTCTCTGGTCCGGGAGTCCGCGCCACCCCGGAGACTACGT 416
OY	728 CCTGTGCGTAGCTTTGGCCCGAGCGTCATCCACTACCGCGGTGTGCAACCGCAGCGCA 787
DB	417 CCTGTGCGTAGCTTTGGCCCGAGCGTCATCCACTACCGCGGTGTGCAACCGCAGCGCA 476
OY	788 CCTGCATCATGTAGAGGCGCGTCTCTTGCAACCTCATGGACATGGAGGACATTACAG 847
DB	477 CCTGCATCATGTAGAGGCGCGTCTCTTGCAACCTCATGGACATGGAGGACATTACAG 536
OY	848 CAAGGACAAAGGCGCTATCTGCACCAACCTGGTAGACCAAAAGCGGAACACGGGACCAA 907
DB	537 CAAGGACAAAGGCGCTATCTGCACCAACCTGGTAGACCAAAAGCGGAACACGGGACCAA 596
OY	908 GTCCGCGCAGAGAGAGCTGGCCAGGCGCGGCTGGTTACTGAACCTGCAAGCTTTTACATT 967
DB	597 GTCCGCGCAGAGAGAGCTGGCCAGGCGCGGCTGGTTACTGAACCTGCAAGCTTTTACATT 656
OY	968 GGGAGACAGATCGGAGAGGAGAGATTGGAGCTTCTGCGAGGCTGAGTACCTGAGGCA 1027
DB	657 GGGAGACAGATCGGAGAGGAGAGATTGGAGCTTCTGCGAGGCTGAGTACCTGAGGCA 716
OY	1028 AAAGGTGGCCGTGAAGAATATCAAGTGTGATGTGACAGCCAGCGCTTCTGTGAGAGAC 1087
DB	717 AAAGGTGGCCGTGAAGAATATCAAGTGTGATGTGACAGCCAGCGCTTCTGTGAGAGAC 776
OY	1088 GGGCGTCATGACGAAGATGCAACACGAACCTGGTGGCTCTCCGGGCGGTGATCTCTCA 1144
DB	777 GGGCGTCATGACGAAGATGCAACACGAACCTGGTGGCTCTCCGGGCGGTGATCTCTCA 836
OY	1148 CCAGGGGCTGTACATTGTTCATGAGACAGTGAAGGCAACCTGGTGAACCTTCTGCG 1207
DB	837 CCAGGGGCTGTACATTGTTCATGAGACAGTGAAGGCAACCTGGTGAACCTTCTGCG 896
OY	1208 GACCCGGGGGTGAAGCCTCTGTGAACACCGCTCAGCTCTGCAAGTTTCTGTGACGTGGC 1267
DB	897 GACCCGGGGGTGAAGCCTCTGTGAACACCGCTCAGCTCTGCAAGTTTCTGTGACGTGGC 956
OY	1268 CGAGGGCATGATGATACCTGGAGAGACAGAACTTGTGACGCGGCACTGGGCGGCCGCA 1327

```

Db 957 CGAGGGCATGAGTACCTGGAGAGCAAGAACCTTGTGACCGGCACTGGCCGCCGCA 1016
Qy 1328 CATCTGTCTCTACAGAGACTGTGTGCCAAGTCAAGCACTTTGGCTGGCCAAAGCCGA 1387
Db 1017 CATCTGTCTCTACAGAGACTGTGTGCCAAGTCAAGCACTTTGGCTGGCCAAAGCCGA 1076
Qy 1388 GCGAGAGGGGCTTACACTCAAGCCGGCTGCTCAAGTGGAGGGCCGCCGAGGCTCTCAA 1447
Db 1077 GCGAGAGGGGCTTACACTCAAGCCGGCTGCTCAAGTGGAGGGCCGCCGAGGCTCTCAA 1136
Qy 1448 ACACGGGAATTCACCAAGTGTGATGTCTGAGTGTGGGTTGGGCTGCTCTGGAGGT 1507
Db 1137 ACACGGGAATTCACCAAGTGTGATGTCTGAGTGTGGGTTGGGCTGCTCTGGAGGT 1196
Qy 1508 CTTCTCATATGACAGGCGCTCCGTAACCTTAATATGCACTGAAGAGGTGTGAGAGCCGT 1567
Db 1197 CTTCTCATATGACAGGCGCTCCGTAACCTTAATATGCACTGAAGAGGTGTGAGAGCCGT 1256
Qy 1568 GGAGAGGGGTACCGCATGGAACCCCGGAGGGCTGTTCAGAGCCCGTGCACGTCCTCAT 1627
Db 1257 GGAGAGGGGTACCGCATGGAACCCCGGAGGGCTGTTCAGAGCCCGTGCACGTCCTCAT 1316
Qy 1628 GAGCAGCTGCTGGAGGAGCAGAGCCCGCCGCGCCACACCTTCCGCAAACTGGCCGAGAA 1687
Db 1317 GAGCAGCTGCTGGAGGAGCAGAGCCCGCCGCGCCACACCTTCCGCAAACTGGCCGAGAA 1376
Qy 1688 GCTGGCCCGGAGAGCTACGACAGTGAAGTGCCTCCGCTCTCAAGGAGAGAGCCGA 1747
Db 1377 GCTGGCCCGGAGAGCTACGACAGTGAAGTGCCTCCGCTCTCAAGGAGAGAGCCGA 1436
Qy 1748 CGGCTCACCTGCGCCGAGGAGCCGCTGACCCGCTGGGCTGGGCGCTTGGGCGCC 1807
Db 1437 CGGCTCACCTGCGCCGAGGAGCCGCTGACCCGCTGGGCTGGGCGCTTGGGCGCC 1496
Qy 1808 AGAGAGCCGAGAGAGTGAAGTGCCTGGGAGGAGCTGACACAGGCCCAAGAGAGGTCC 1867
Db 1497 AGAGAGCCGAGAGAGTGAAGTGCCTGGGAGGAGCTGACACAGGCCCAAGAGAGGTCC 1556
Qy 1868 AGGCGGCAAGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1927
Db 1557 AGGCGGCAAGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1616
Qy 1928 GCGGCGCGTGGACACCCGACCTGCGAAGATGATGCGCGCTGAAGAGGATTTCTAA 1987
Db 1617 GCGGCGCGTGGACACCCGACCTGCGAAGATGATGCGCGCTGAAGAGGATTTCTAA 1676
Qy 1988 GGAAGCTTAAAAA 2000
Db 1677 GGAAAAA 1689

```

RESULT 6
AAZ29701
ID AAZ29701 standard; cDNA; 2187 BP.
XX
AC AAZ29701;
XX
DT 22-MAR-2000 (first entry)
XX
DE wild-type human c-Src tyrosine kinase cDNA.
XX
KW Angiogenesis; tyrosine kinase; Src; inhibition; activation; modulate;
KW human; viral expression vector; replication competent; mutant Src;
KW inflammatory disease; arthritis; rheumatoid arthritis; restenosis;
KW diabetic retinopathy; osteoporosis; cancer; ss.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT 134..1486
FT CDS
FT /*tag= a
FT /product= "Human c-Src tyrosine kinase"
FT /note= "Src used to modulate angiogenesis"

```

XX XX
PN W09961590-A1.
XX XX
PD 02-DEC-1999.
XX XX
PF 28-MAY-1999; 99W0-US11780.
XX XX
PR 29-MAY-1998; 98U5-0087220.
XX XX
PA (SCRI ) SCRIPPS RES INST.
XX XX
PI Chersesh DA, Elliceiri B, Schwartzberg PL;
XX XX
DR WPI: 2000-116335/10.
XX XX
PT P-PSDB; AAY44448.
XX XX
PS Using tyrosine kinase Src for modulating angiogenesis in tissues useful
XX XX in, e.g. treatment of chronic articular rheumatism -
XX XX
PS Claim 1; Fig 3; 80pp; English.
XX XX
CC The present sequence is the cDNA, encoding the wild-type human c-Src
CC tyrosine kinase. This sequence encoding the Src protein can be used to
CC modulate angiogenesis. When the Src protein is inactivated, angiogenesis
CC is inhibited, while when it is activated, angiogenesis is potentiated.
CC The modified or variant Src can be used to treat inflammatory diseases
CC like, arthritis, rheumatoid arthritis, diabetic retinopathy, restenosis,
CC osteoporosis and cancer associated disorders.
XX XX
SO Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other;

```

Query Match 27.4%; Score 547.6; DB 21; Length 2187;
Best Local Similarity 64.9%; Pred. No. 2,4e-100;
Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

```

Qy 395 CTGGGCCCCGGGACCCAGTATACCAATGTGAGACACCCGCCCAAGCCAGGGA 454
Db 154 CTGGGCCCCGGGACCCAGTATACCAATGTGAGACACCCGCCCAAGCCAGGGA 213
Qy 455 GCTGGCTTCCGAGAGGCGAGCTGTGACCATCTGAGAGCCCTGGCAAGAGCTG 514
Db 214 CTGGCTTCCGAGAGGAGACCGTGTGACCATCTGAGAGCCCTGGCAAGAGCTG 273
Qy 515 GTACCCGCTAAGACACCAACAGTGAAGAGAGGCTCTGGACCTGGGCGCTGCG 574
Db 274 GTACCAAGGCA---AAACAGAGTGGCCCTGAGGAGCATCCAGCAACTACGTCA 330
Qy 575 GAGCGGGAGGCGCTCTCCGACAGCCCAAGCTCAGCTCAGCTGCTGCTGCTGCTG 634
Db 331 GAGCGGGAGGCGCTCTCCGACAGCCCAAGCTCAGCTCAGCTGCTGCTGCTGCTG 390
Qy 635 GATCTGGGCGAGAGAGCTGTCCAGCTGACAGCTCCCGAGAGTGGCTGTCTGCTG 694
Db 391 GATCTGGGCGAGAGAGCTGTCCAGCTGACAGCTCCCGAGAGTGGCTGTCTGCTG 450
Qy 695 GCGGAGAGTCCGCGCCGACCCCGGACACTGACGTCTGTGCTGAGTCTTGGCCGAGCT 754
Db 451 GCGGAGAGAGTCCGCGCCGACCCCGGACACTGACGTCTGTGCTGAGTCTTGGCCGAGCT 510
Qy 755 CATCTCACTACCGGTGTGACAGCGGAGCCGACCTCAACAAATGATGAGGCGCTGCTT 814
Db 511 GAGCAGCTACCGGTGTGACAGCGGAGCCGACCTCAACAAATGATGAGGCGCTGCTT 570
Qy 815 CTGCAACTCATGATGATGTGAGGATTTACAGCAAGAGAGGCGCTATCTGACCAA 874
Db 571 TGAGAACTCATGATGATGTGAGGATTTACAGCAAGAGAGGCGCTATCTGACCGG 630
Qy 875 GCTGTGTGAGACCAAGAGGAGCAAGCAAGTGTGCGGAGAGAGAGTGGCCAGGC 934
Db 631 CCTCATTTAAACCAAGAGTGTGAGGAGGAGGAGGCGCCAGAGATGATTTTACCGCAG 690
Qy 935 GCGCTGTATCACTCAAGCTCAGCATTTGACATTTGGAGAGCAGAGATGAGAGGAGATT 994

```

[illegible]

Query Match	Best Local Similarity	Matches	Score	DB	Length
27.4%;	64.9%;	846;	547.6;	22;	2187;
		Conservative		Pred. No. 2.4e-100;	
		0;	Mismatches 449;	Indels 9;	Gaps 2
35	CTGGGCCCCGGGACCCAGTGTATCTACCAATGCGAGACACCCGCCCAACGAGGGA	454			
154	CTGGGCAATCCGGTATACGAATGTATTGGCCAAAGTACAACTTCCACGGGACACTGCCGAGCAGA	213			
455	GCTGCGCTTCCGACGAGGGGAGCTGTGTCACATCTCTGGAGGCGCTCGAAGAACAGCTG	514			
214	CCTGCGCTTCTGCAAAAGGAGAGCTGTCTACCACTTGTGGCGGCTACCAAGGAGCCCAACTG	273			
515	GTAACCGCTCAAGACACCAACACGTTGAGACAGGAGGGGCTGTGTCAGACTGGGGCGCTGCG	574			
274	GATCAAAAGCCA--AAACAAGAGTGGGGCGGTGAGGGGACATCATCCAGCCCAACTACGTCCA	330			
575	GGAGGGGAGAGGCGCTTCCGACAGACCCCAAGTGTGACGCTCATGCGTGGTTCCAGGGAA	634			
331	GAGGCGGAGGGGCTGAAAGGCGGATACCAACTCAAGCTTCTTGATTCACGGCAA	390			
635	GATCTCGGAGCCAGGAGGCTGTCTACAGCTGTGACGCTCCCGAGATGGGCTTCTCTGCT	694			
391	GATCAACAGGAGGAGCGGTGAACGCGGCTTGTACCGCGCGGAGACAGCGCTGTCTGTG	450			
695	GGGGAAGTCCGGCGCCACACCCGGCGAGCTACAGTCTCTGTGCGTGAAGCTTTGGCCGAGACT	754			
451	GGGGAGAGCACCAACTACCCCGGAGACATCAACGCTGTTCGTGAGCTGTGAGCGGCAAGGT	510			


```

OY 755 CATCCACTACCGCTGCTGCACCGCGGACCCACCTACAAATGCATGAGCCGCTGTTCTT 814
DB 511 GGAGCACTACCGCATCATGTACATGCCAGCACTGAGCTACGACGAGGAGTGACTT 570
OY 815 CTGCACCTCATGACATGGTGGAGCATTTACAGCAAGACAGGCGCTATCTCCACCA 874
DB 571 TGAACACTCATGACATGGTGGAGCATTTACAGCAAGACAGGCGCTATCTCCACCA 630
OY 875 GCTGGTGAACCAAGGCGAAGAACACGAGCCCAAGTCCGCGGAGAGAGAGTCCGACGAG 934
DB 631 CTTCTATTAACCAAGGCGAAGAACACGAGCCCAAGTCCGCGGAGAGAGTCCGACGAG 690
OY 935 GGGCTGTTACTGACACTGCAGCATTTGACATTTGGAGCACAGATCGGAGGAGAGATT 994
DB 691 CGGCTGGGCGCTGAAACATGGAAGGAGTGAAGCTGCTCAGACACATCGGAGGAGATT 750
OY 995 TGGAGCTGCTCTGCAGAGGTAGTACTGAGGCAAAAGGTGGGCAATATCAATG 1054
DB 751 CGGAGAGCTGATGCTGGGCGATTAACGAGGAGAACAAAGTCCGCTCAAGTGCATTAGAA 810
OY 1055 TGATGTGACAGCCCGCTCTCTGACAGAGAGCGCGCTATGACGAAGATGCAACAGA 1114
DB 811 CGAGCGCACTGCCAGGCGCTCTCTGAGGAGCGCTCAGTATGACGACACTCGCGGAGTAG 870
OY 1115 GAACCTGCTGCTCTCTGAGGCGATCCCTGCACACAG-----GGCTGTACATTTGTCAT 1168
DB 871 CAACCTGCTGCTCTCTGAGGCGATCCCTGAGGAGAGAGAGGCGGCTCAATCGTCAC 930
OY 1169 GAGAGCACTGAGCAAGGCGCAACTGTGTAATTTCTGCGAGACCGGCGGCTGAGCCCTCGT 1228
DB 931 TGAGTACATGCGCAAGGCGAGGCGCTGTGAGACTACCTCGGCTTAAGGGGTGGCTAGTGCT 990
OY 1229 GAACACCGCTCAGCTCTCTGCAAGTTTCTCTGCACTGCGGCGAGGCGATGAGTACTGA 1288
DB 991 GGGCGGAGACTCTCTCTCTCAAGTTTCTCTGATAGTCTGCGAGGAGGCAATGAAATACCTGGA 1050
OY 1289 GAGCAAGAGCTGTGACAGCGGAGCGGCGCGCAATCTGCTCAAGAGACT 1348
DB 1051 GGGCAACAAATTTCTGATCATGAGACCTGCGTCCCGCAATGCTGCTGAGAGCA 1110
OY 1349 GGTGGCCAAAGTCAAGGACTTTGCGCTGCGCAAGCGGAGCGGCTAGACTCAAG 1408
DB 1111 CGTGGCCAAAGTCAAGGACTTTGCTCAACGAAGGCGCTCAAGACCCAGAGACAGG 1170
OY 1409 CCGCGTCCCGCTCAATGAGAGCGCGCGCAAGGCTCTCAAAACAGGGAAGTTCAACAGCAA 1468
DB 1171 CAAGCTGCAAGTCAATGAGACAGCCCTGAGGCGCTGAGAGAGAAATTTCTCCACTAA 1230
OY 1469 GTGCGATGTCTGAGATTTTGGGCTGCTGCTGAGAGGCTTCTCATATGAGAGGCGCTCC 1528
DB 1231 GTCTGACGTGTGAGATTTTGGGCTGCTGCTGAGAAATCTACTCTTGGGCGAGTGGC 1290
OY 1529 GTACCTTAATATGCTCAAGAAAGGTGTGCGAGGCGGTGAGAGAGGAGTCCGATGGA 1588
DB 1291 TTATCCAAATTTCCCTGAGAGAGCTGCTCCCTCGGTGAGAGAGGAGTCAAGATGGA 1350
OY 1589 ACCCGCGGAGGCTGTGCAAGCGCGCGCTGACAGCTCATAGAGAGTCTGCGAGGAGCA 1648
DB 1351 TGCCCGCGAGGCTGTGCGCGCGCACTGTATGAAGATGAAGAACTGTGCGCACTGGA 1410
OY 1649 GCCCGCGCGCGCGCACTGCGCAAACTGGCGGAGAGAGCTGG 1692
DB 1411 CGCGCGCATGCGGCGCTCTCTCTTCTACAGCTCCGAGAGCAAGCTTG 1454

```

```

RESULT 8
ABK84602
ID ABK84602 standard: cDNA: 2187 BP.
AC
XX ABK84602:
XX
DT 14-AUG-2002 (first entry)

```

```

XX XX Human CDNA differentially expressed in granulocytic cells #1173.
DE Human; ss; granulocytic cell; DNA chip; bacterial infection;
XX viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; peritonitis;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-Apr-2002.
XX
PE 03-Oct-2001; 2001WO-US30821.
XX
PR 03-Oct-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI: 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity.
XX
XX Claim 1; SEQ ID No 1173; 114pp; English.
XX
PS The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) Gs by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a subject having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GCA; M3 is useful for screening an agent capable of modulating
CC GCA preferentially in an inflammation (in a tissue); M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC peritonitis; also bacterial infection, viral infection, and
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WPI at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other;

```


Query Match 27.4%; Score 547.6; DB 24; Length 2187;
 Best Local Similarity 64.9%; Pred. No. 2.4e-100;
 Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

```

QY 395 CTGGGCCCCGGGCAACAGTGTATCACAAATGGAGACACCCGCCCAAGCCAGGGGA 454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 154 CTGGCCATCCGCTAGAAATGTATTGGCAATACAACTCCACGCACTGGCGAGACAGA 213
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 455 GCTGGCCTTCGGCAAGGGGACGTGTGTACCTCTGGAGGCTTGGCAAGAGAGT 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 CTTGCTCTTGGCAAGGAGACGTGTACCTCTGGAGGCTTGGCAAGAGAGT 273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 515 GTACCGGCTCAAGACACCAAGGTGACAGAGAGGGGCTGTGGAGCTGGGGGCGT 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 274 GTACAAAGCCA---AAAACAAGGTGGGCGGTGAGGGGCTATCCAGCAACTACTTCA 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 575 GGAGCGGGAGGCGCTTCCGAGACCCCAAGCTCAGCTCATGGCGTTCACAGGGAA 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 331 GAAGCGGGAGGGCGTGAAGGGCGGTACCAAACTCAGCTCATGGCGTTCACAGGGCA 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 635 GATCTCGGGCCAGAGGCTGTCCAGACAGCTCAGCTTCCGAGATGGGCTGTTCGT 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 391 GATCACAAGGGAGAGGCTGAGCGGCTTCTTACCCGCGGAGACAGGCTGTTCGT 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 695 GCGGGAGTCCGGGCGCCACCCCGGACTAGTCTGTGGTGGAGCTTGGCGCGGAGCT 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 451 GCGGGAGAGCACCAACTACCCCGGAGCTACAGCTGTGGCTGAGCTCGAGCGGAGGT 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 755 CATCCACTACCGGCTGTGTGACCGGAGCGGCGACCTCAATCATGTGAGGCGTGTTC 814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 511 GGAGCAGCTACCGCATCATGTATGACATGCGAGAGTGTGATGAGAGGAGGTACTT 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 815 CTGCAACTCATGTGACATGTGTGACATTTACAGCAAGACAAAGGCGCTTATCTGACCAA 874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 571 TGAGAACTCATGTGAGCTGTGAGACTACACTCAAGCAGATGTGACTGTATCCGC 630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 875 GCTGTGAGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 631 CTTCTATTAAACCAAGGATCATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 935 GGGCTGTGTACTGAACCTGTGACATTTGACATTTGGAGACACAGATCGAGAGGAGAGT 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 691 CGGCTGGGCGCTGAACATGAGAGAGCTGAAGCTGTGACAGCATCGGAGAGGGAGATT 750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 995 TGGAGCTGTCTCTGAGGCTGTGACATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 751 CGGAGAGCTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 810
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1055 TGATGTGACAGGCGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 811 CGAGGCCAGCTGCCAGGCGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 870
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1115 GAACCTGTGTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 871 CAACCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 930
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1169 GGAGCAGCTGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 931 TGAGTACATGTGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 990
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1229 GAACACGCTCAGCTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 991 GGGGAGAGACTGTCTCTCAAGTGTCTGCTAGATGTCTGCAAGGCGCAAGGAGGAGT 1050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1289 GAGCAGAAGCTTGTGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1051 GGGCACAATTTTGTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1349 GGTGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1111 CGTGGCCAAGGTCAGCGACTTGTGCTTACCAAGAGGAGGAGGAGGAGGAGGAGGAGT 1170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 1409 CCGGCTGCCCTCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1171 CAAGCTGGCAGTCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1469 GTGGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1231 GTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1529 GTACCTTAAATGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1291 TTATCCAAAGATTTCCCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1589 ACCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1351 TGCCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1649 GCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1411 GCGCGCATGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 1454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
ABL62918
ID ABL62918 standard; DNA; 2187 BP.
XX
XX ABL62918;
XX
XX AC
XX
XX 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:1255.
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumor; cancerous;
XX cytostatic; gene therapy; antineoplastic; Wilms tumor; adenocarcinoma;
XX gene; ds.
XX
OS Homo sapiens.
XX
XX WO200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
XX
XX 05-JUN-2000; 2000US-209473P.
XX
XX 05-JUN-2000; 2000US-209531P.
XX
XX 18-SEP-2000; 2000US-233133P.
XX
XX 18-SEP-2000; 2000US-233617P.
XX
XX 20-SEP-2000; 2000US-234009P.
XX
XX 20-SEP-2000; 2000US-234034P.
XX
XX 20-SEP-2000; 2000US-234052P.
XX
XX 20-SEP-2000; 2000US-234509P.
XX
XX 22-SEP-2000; 2000US-234523P.
XX
XX 22-SEP-2000; 2000US-234567P.
XX
XX 25-SEP-2000; 2000US-234923P.
XX
XX 25-SEP-2000; 2000US-234924P.
XX
XX 25-SEP-2000; 2000US-235072P.
XX
XX 25-SEP-2000; 2000US-235082P.
XX
XX 25-SEP-2000; 2000US-235134P.
XX
XX 25-SEP-2000; 2000US-235808P.
XX
XX 26-SEP-2000; 2000US-235637P.
XX
XX 26-SEP-2000; 2000US-235638P.
XX
XX 27-SEP-2000; 2000US-235711P.
XX
XX 27-SEP-2000; 2000US-235720P.
XX
XX 27-SEP-2000; 2000US-235840P.
XX
XX 27-SEP-2000; 2000US-235863P.
XX
XX 28-SEP-2000; 2000US-236028P.
XX
XX 28-SEP-2000; 2000US-236032P.
XX
XX 28-SEP-2000; 2000US-236033P.
XX
XX 28-SEP-2000; 2000US-236034P.
XX
XX 28-SEP-2000; 2000US-236034P.
XX
XX 28-SEP-2000; 2000US-236109P.
XX
XX 28-SEP-2000; 2000US-236111P.
XX
XX 29-SEP-2000; 2000US-236842P.
XX
XX 29-SEP-2000; 2000US-236891P.

```

PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237596P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppe DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 PS Claim 1; SEQ ID 1255; 44pp; English.
 XX
 XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in AB161664
 CC to AB170110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC adenocarcinoma, carcinoma, kidney, prostate or pancreatic cancer,
 CC infiltrating lobular cancer, clear cell cancer, infiltrating ductal cancer,
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 CC
 SQ Sequence 2187 BP; 460 A; 651 C; 648 G; 428 T; 0 other;
 Query Match 27.4%; Score 547.6; DB 24; Length 2187;
 Best Local Similarity 64.9%; Pred. No. 2.4e-100;
 Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

DB 451 GCGGAGAGACCACTACCCCGGAGACTACAGCTGTGCTGAGCGGCAAGT 510
 QY 755 CATCACTACCGGTGCTGACCGGACGCGCACTCACTCATGATGAGCCGTGTTT 814
 DB 511 GAGACATACCGCATCATGTACCATGACGACCAAGCTGAGATGAGAGGTTACTT 570
 QY 815 CTGCACCTGATGACATGTGTGAGCATTTACAGCAAGAGAGGCGCTATCTGACCAA 874
 DB 571 TGAAACCTCATGCTGCTGAGGACACTACCTCAGACGCGAGTGTGACGCG 630
 QY 875 GCTGTGAGACCAAGGGAAGAACAGGAGCAAGTGGCCGAGAGAGTGGCCAGGCG 934
 DB 631 CTTCTATTAACCAAGGTCTATGAGGACAGTGGCGCCGAGAGTGTTCACGCGAG 690
 QY 935 GGGCTGTTTACTGACCTGACGATTTGACATTTGGAGCAGACAGATGAGAGGAGATT 994
 DB 691 CGGCTGGGCGCTGACATGAAAGAGAGTGAAGCTGTCTCAACCATCGGAGGAGATT 750
 QY 995 TGAAGCTGTCTGACAGGCTGAGTACCTGGGCAAAAGCTGGCCGTGAAGATATCAAGTG 1054
 DB 751 CGAGAGAGTGTATCTGGGCGATTACGAGGGAACAAAGTGGCGTCAAGTGAATTAAGAA 810
 QY 1055 TGATGTGACAGCCGAGGCTTCTGAGACGAGGCGGCTATGCAAGATGCAACGAA 1114
 DB 811 CGAGCCACTGCGAGGCTTCTGAGGCTGAGGCTGATGACGACTGGCGCTAG 870
 QY 1115 GAACCTGTGCTGCTCTGAGGCGTATCTGACAGAG-----GGCTGTACATTTGTAT 1168
 DB 871 CAACCTGTGCTGCTCTGAGGCGTATCTGAGGAGAGAGGAGGCTGTACATCTGCTAC 930
 QY 1169 GAGACGCTGAGCAAGGCAAGCTGTGAACTTTCTGCGAGCCGCGGCTGAGACCTCTCT 1228
 DB 931 TGAGTACATGCAAGGAGGAGGCTTGTGAGCTACCTGCGCTGAGGAGTGTGCTGCT 990
 QY 1229 GAACACGCTGCTCTCTGAGGCTTGTGAGCTGAGGAGGAGGAGTGTGAGTACCTGGA 1288
 DB 991 GGGCGGAGAGTGTCTCTGAGGCTTGTGAGCTGAGGAGGAGGAGTGTGAGTACCTGGA 1050
 QY 1289 GAGCAGAGAGCTTGTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1348
 DB 1051 GGGCAGCAGAGTGTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1110
 QY 1349 GGTGGCCAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1408
 DB 1111 GGTGGCCAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1170
 QY 1409 CCGGCTGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1468
 DB 1171 CAGGCTGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1230
 QY 1469 GTGCGATGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1528
 DB 1231 GTCTGAGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1290
 QY 1529 GTACCTTAATATGTCACTGAAAGAGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1588
 DB 1291 TTATCCAGAGATTTCTCTGAAAGAGTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1350
 QY 1589 ACCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1648
 DB 1351 TGCCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1410
 QY 1649 GCGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1692
 DB 1411 CGCCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1454
 RESULT 10
 AA246489
 ID AA246489 standard; DNA: 2420 BP.
 XX
 AC AA246489;
 XX

DT 13-MAR-2000 (first entry)

XX

DE PKA substrate; Csk-family protein encoding DNA.

KW Protein kinase A; PKA; PKA signaling pathway; phosphorylation; cancer;

RN kinase substrate; immunosuppressive disorder; proliferative disease;

CC HIV infection; AIDS; immunodeficiency; autoimmune disease; Chk; Lsk; Hyl;

KM systemic lupus erythematosus; Csk-family; Matk; Clk; Bnk; Ntk; Ss.

XX

OS Homo sapiens.

FH

FT Key Location/Qualifiers
CDS 413..1765
FT /*tag= a

PX MO9962315-A2.

PN

PD 02-DEC-1999.

XE

PF 27-MAY-1999; 99WO-GB01680.

PR 27-MAY-1998; 98NO-0002419.
PT 30-DEC-1998; 98US-0114240.

PA (LAUR-) LAURAS AS.
PI (JONE/) JONES E L.

DZ Hansson V., Levy FO, Mustelin T, Skalhogg BS, Sundvold V, Tasken K;
DR Wang T, Altman A, Munshi A;
PS WPI: 2000-086801/07.
PP P-PSDB: MAY49418.

PT Altering the activity of protein kinase signalling pathways, used for treating immunosupressive disorders, e.g. AIDS, proliferative disorders, e.g. cancers or autoimmune diseases -

XX

PC Claim 11; Page 87-89; 11pp: English.

The invention provides a novel method of altering the activity of the protein kinase A (PKA) signalling pathway in a cell that comprises altering the extent of phosphorylation of one or more PKA substrates, or kinase substrates downstream in the PKA signalling pathway. Pharmaceutical compositions containing a nucleic acid molecule that encodes a PKA substrate, or fragment, precursor or functionally equivalent variant, where the sequence is modified to alter its susceptibility to phosphorylation by PKA can be used for treating a disorder exhibiting abnormal PKA signaling activity, immunosuppressive disorders or proliferative diseases. They can be used for treating e.g. HIV infection, AIDS, common variable immunodeficiency or cancers. Conditions in which upregulation of the PKA pathway is required, such as autoimmune disease, e.g. systemic lupus erythematosus, may also be treated. The present sequence represents a DNA sequence encoding a PKA substrate, wherein the substrate is in the Csk-family, preferably Csk, Chk, Lsk, Hyl, Matk, Clk, Bnk or Ntk.

SQ Sequence 2420 BP; 471 A; 771 C; 703 G; 475 T; 0 other:

Query Match 27.4%; Score 547.6; DB 21; Length 2420;
Best Local Similarity 64.9%; Pred. No. 2.5e+100;
Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2

OY CTGGGCCCCCGGCACGCCAGTGTATCATCCAATAAGCGAGCACACC CGGCCCAACGAGA 454
| | | | | | | | | | | | | | | | | | | | |
Db 433 CTGGCCATTCCGGTGATACAAGATGTATTCCTGAAGTAACAATTCCACGCGACTCCGAGCAGGA 492
| | | | | | | | | | | | | | | | | | | | |
OY 455 GC TGAGCCTTTCCGCAAGGGCGAGCGTAGTGTCACCATCTCGAGGCGCTCGAGAACAGAAGCTG 514
| | | | | | | | | | | | | | | | | | | | |
Db 493 CCTGCCTTTCTGCAAAGGAGAGAGCTCTCACCATTTGTGGCGCTGACCAAGAGACCCAACTG 552
| | | | | | | | | | | | | | | | | | | | |
OY 515 GTACCGCTCAAGACACACACAGTGAGACAGAGAGGGGCTGCTGGCAGCTGGGGGCGCTGGG 574
| | | | | | | | | | | | | | | | | | | | |
Db 553 GTAACAAAACCC--AAAAACAAGGTGGGCGTGAGAGGCAATCATCCCACGCAACTAGCTCA 609

QY	575	GGACGGGAGGCCCCCTCCGAGACGCCCAAGCTCAAGCTCATGCGCGTGGTTCCACGGGAA	634
Db	610	GAAGCGGAGGGGCTGAAAGCGCGGTACCAAACTAGCTCTATGCTTGTTCCACGGCAA	669
QY	635	GATCTCGGGCCAGAGGCTGTCCAGACGTGCAGCCCTCCGAGAGTGGCTGTTCCTGT	694
QY	755	CATCAGTTACCGCGCTGTGTGCACCGCGAGCGCACCTCACAATCATGAGCGCGTCTT	814
Db	790	GGAGACTTACCGCATATGATACCATGCGCAGCAAGCTAGCATGCAAGAGAGGTACTT	849
QY	815	CTGCACCTCATGAGACATGTTGGAGCTATACGACAGAGCAAGGCGCTATCTGCACAA	874
Db	850	TGAAGACCTCATGACGTGTGTGGAGCACTACCTCAGACGCAAGATGACTCTGACGC	909
QY	875	GCTGTGAGACCAAGGGGAAACACGGGACCAAGTCGGCCGAGAGGAGAGCTGGCGCAAGCG	934
Db	910	CCTATTAAACCAAAAGTCAATGGAGGACAGTGGCGCCGCCAGATAGTTCTTACCCAC	969
QY	935	GGCGTGTCTTACGAACTTCGACGATTTTGACATTGGGAGCAGACAGATCGAGAGAGATT	994
Db	970	CGCGTGGGCCCTGAAACATGAAGAGCTGAAAGCTCTGCAGACATCGGAAAGGGGAGTT	1029
QY	995	TGAGAGCTGCTCGTAGGGTAGTACCTGCGGGCAAAAGGTGGCCGTGAAGAATATCAAGTG	1054
Db	1030	CGGAGAGCTGATGTGTGGGCAATTACCAAGGGAACAAAGTGGCCCTCAAGTGCATATGAAG	1089
QY	1055	TGATGTACAGCCAGGCGCTTCTCGAGACGAGCGCGCTATGACGAAGATGCAACAGA	1114
Db	1090	CGAGCGCACTGCCAGGCGCTTCTCGGTGAAAGCTCATGTATGACGCAACTCGCGCATAG	1149
QY	1115	GAACCTGTGCGTCTCTGGGCGTGATCTGCACAG-----GGCTGTACATTGTCAAT	1168
Db	1150	CAACCTGTGTAGCTCTCTGGGCGTGATCTGTGAGAGAAAGGCGCGCTCAATCACTGCAC	1209
QY	1169	GGACGACGTGAGCAAGGGCAACCTGGTGAATTTCTGCGGACCCGGGGGTGAGCGCCGT	1228
Db	1210	TGAGTACATGCGCCAGAGGAGACCTTGTGTGACACTCTCGGCTTAGAGGGTGGTATGCT	1269
QY	1229	GAACACCGCTACGCTCCCTGAGTTTCTCTGCACGTGGCGGAGGCGATGAGTACCTGA	1288
Db	1270	GGGGGAGACGTGTCTCTCAAGTTCTCGTAGATGTCTGCGAGGCGATGGAATCTGGA	1329
QY	1289	GAGCAABAAGCTTGTGCACCGGACCTGGCGCCCGGCAACATCTGCTTCAGAGACT	1348
Db	1330	GGGCAACAATTTCTGTGATCGAGACGCTGTGCGCCGCAATGTGCTGTGTGTGGACAA	1389
QY	1349	GGTGGCCAAAGGTACGCACTTTGGCTCGGCAAAAGCGAGCGAAGAGGGGTAGACTAAG	1408
Db	1390	CGTGGCCAAAGTCAAGCACTTTGTGTCTACCAAGAGAGCGCTCAGACACCCAGGACAGGG	1449
QY	1409	CCGGCTGCGGCTCAAGTGAAGCGCGCCGAGGCTCTAACAACGCGGAATTCACACAA	1468
Db	1450	CAACCTCCCAAGTCAAGTGAAGACGCCCTCGAGGCCCTCAGAGAGAAATTTCTCCATAA	1509
QY	1469	GTCGAGTGTGGAAGTTTGGGGGTGTGACCTGGAGAGGTCTTCATATGAGACGGGCTC	1528
Db	1510	GTCGACGTGTGGAGTTTCGGAATCTTCTCTGTGGAAATCTACTCTTTTGGCGAGTGC	1569
QY	1529	GTACCTTAAATGTCACTGAAAGAGTGTGCGAGGCCGTGAGAAAGGGGTACCGCATGA	1588
Db	1570	TTATCCAAAGATTCCTCTGAAGAGCGTCTCCCTCGGGGTGAGCAAGGCTTCAAGATGA	1629
QY	1589	ACCCCGGAGGGGTGTCCAGGGCCCGGACACGTCTCATAGCAGCGTGTGGAGCGAGA	1648
Db	1630	TGCCCCGAGCGGCTGCCCCCGAGTCTATGAAGTATGAAGACTGTCTGGACACTGGA	1689

QY 1649 GCCGCGCCGCGGACCCCTTCGCAAACTGCGCCGAGAGCTGG 1692
DB 1690 CGCGCGCATGCGCCCTCTCTACAGCTCGAGAGCAGCTTG 1733

RESULT 11
ID ACC50120 standard; cDNA; 2420 BP.
XX
AC ACC50120;
XX
DT 12-JUN-2003 (first entry)
XX
DE Breast cancer associated cDNA sequence SEQ ID NO:87.
XX
KW Human; breast cancer; cytostatic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200304989-A2.
XX
PD 16-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US19669.
XX
PR 21-JUN-2001; 2001US-299887P.
PR 27-JUN-2001; 2001US-301572P.
PR 18-JUL-2001; 2001US-306501P.
PR 25-SEP-2001; 2001US-325002P.
PR 05-MAR-2002; 2002US-362585P.
PR 14-MAY-2002; 2002US-380391P.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Lallian JE, Ganavarapu M, Glatz K, Hoersch S, Kamatkar S, Mertens M;
PI Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE, Bast RC;
PI Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
PI WPI; 2003-210381/20.
DR P-PSDB; ABR47428.
XX
XX
XX Breast cancer diagnosis or treatment by comparing the level of
XX expression of a marker in a patient sample with that in the control
XX non-breast cancer sample -
XX
XX Claim 1; SEQ ID 87; 128bp; English.
XX
XX The present invention describes a method for assessing whether a patient
XX is afflicted with breast cancer. The method comprises comparing the level
XX of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
XX ABR47386 to ABR47632) in a patient sample and the normal level of
XX expression of the marker in a control non-breast cancer sample, where a
XX significant increase in the level of expression of the marker in the
XX patient sample and the normal level is an indication that the patient is
XX afflicted with breast cancer. The breast cancer associated sequences
XX from the present invention have cytostatic activities and can be used in
XX gene therapy. The method is useful for diagnosing and treating breast
XX cancer.
XX
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2420 BP; 471 A; 771 C; 703 G; 475 T; 0 other;

Query Match 27.4%; Score 547.6; DB 25; Length 2420;
Best Local Similarity 64.9%; Pred. No. 2,5e-100;
Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

QY 395 CTGGGCGCCGCGGACCCAGTGTATCACAATGCGAGACACCGCCCAAGCCAGGGA 454
DB 433 CTGGGCGATCCGATACGAATGTATGCGCAAGTACCACTCCAGCGACTCCGAGCAGGA 492
QY 455 GCTGGGCTTCCGCAAGGCGACGCTGTACACATCTCTGAGGCTCTGCAAGAACAGAGCTG 514

DB 493 CTTGCCCTTGTGCAAAAGAACGTCACCATTTGTGGCCGTCAACCAAGACCCCACTG 552
QY 515 GTACCGGCTCAAGACACACACGATGACAGAGAGGCTGTGCACTGTGGGGCGCTGGC 574
DB 553 GTACCAAGCCA---AAACAAGGTGGCCCTGTAGGGCATATCCACCAACTATCTCCA 609
QY 575 GGAACGGGAGGACCTCTCCGAGACCCCAAGCTCAGCTCATGCGGTGTTCACGGGAA 634
DB 610 GAAGCGGAGGCGGTGAAGCGGGGTACCAACTAGCTCATGCTGTGGTTCACGGCAA 669
QY 635 GATCTGGGCGCAGAGGCTGTCCAGCACTGTCAAGCTTCCGAGATGGGTGTCTCTGT 694
DB 670 GATTCACAGGGAGAGGCTAGCGCTGTGTACCCCGGAGACAGGCTGTCTCTGT 729
QY 695 GCGGAGTCCGCGCGCCAGCCCGGAGCTACGTCCTGTGCTGAGTGGCCGAGCT 754
DB 730 GCGGGAGAGACCAACTACCCCGGAGACTACAGCTGTGGGTAGCTGCGACGGCAAGT 789
QY 755 CATCCACTACCGGCTGTGCAACGCGAGCGCCACTCAATCATGATGAGCCGTCTT 814
DB 790 GGAGCACTACCGATCATGTATCATATGCCACCAAGCTCAGCATCGACGAGAGGTCTT 849
QY 815 CTGCAACTCATGACATGTGTGAGCATTTACAGCAAGGACAGAGGCGCTATCTGACCAA 874
DB 850 TGAGAACCTCATGACGTGTGTGAGCACTACCTCAGACGACAGATGACTGTGTACGCG 909
QY 875 GCTGTGAGACCAAGGAGGAAACAGGAGCAAGTCCGAGAGAGGCTGGCCAGGCG 934
DB 910 CTCATTTAAACCAAGGTTCATGAGAGGACAGTGGCGCCAGATGATGTTTACCCAG 969
QY 935 GGGCTGTTACTGAACTGTGACGATTTGACATTTGGGAGCAAGATCGAGAGGAGATT 994
DB 970 CGGCTGGGCGCTGACATGAAAGAGAGCTGAGCTGCTCAGACCATCGGAGGGGAGTT 1029
QY 995 TGAAGCTGTCTGTGAGGCTAGTACCTGTGGGCAAAAGTGGCCGTGAAGATATCAAGTG 1054
DB 1030 CGGAGACGTATGCTGGGCAATTACGAGGGAACAAAGTCCGCTCAAGTCAATTAAGAA 1089
QY 1055 TGAATGTACAGCCAGCCCTTCTGAGAGAGAGCGCCGTATGACGAAGATCAACAGCA 1114
DB 1090 CGAGCGCAGTCCAGGCGCTTCTGAGTACCTGAGTACGATGATGACGACGCGGCAATG 1149
QY 1115 GAACCTGTGCTGCTCTGGGCGGTGATCTGCACAG-----GGCTGTACATTTGTCAT 1168
DB 1150 CAACCTGTGCTGCTCTGGGCGGTGATCTGCACAG-----GGCTGTACATTTGTCAT 1209
QY 1169 GGAGCAGTGTGAGAGGAGGCAACCTGTGAGACTTCTGCGGACCCGGGCTGTGAGCCCTGT 1228
DB 1210 TGAATCATGTGCGCAAGGGAGCCCTGTGTGACTACCTGCTGAGGGGTGCTGAGTGT 1269
QY 1229 GAACACCGCTCAGCTCTGCAAGTTTCTGTGCAAGTGGCGAGGCGATGAGTACTGGA 1288
DB 1270 GGGGCGAGACTGTCTCTCAAGTTCTGTGATGTGTGAGAGCCATGGAATATACCTGGA 1329
QY 1289 GAGCAGAGACTGTGTGACCGGCACTGCGCCCGCCGCAACATCTGTCTCAGAGACCT 1348
DB 1330 GGGCAGCAAAATTTGTGATGTGAGAGACCTGTGCGCCGCAAAATGTGTGTCTGAGGCAA 1389
QY 1349 GGTGGCCAAAGTGTGAGGACTTGTGGCCGTGGCAAAAGCCGAGCGGAGAGGGCTGTACATCAAG 1408
DB 1390 GGTGGCCAAAGTGTGAGGACTTGTGTCTCACCAAGAGGCGCTCCAGCACCCAGAGCAGGG 1449
QY 1409 CCGGCTGCCCTGAAGTGTGAGCGGCGCGAGGCTCTCAAAACAGCGGAAGTTTCAACAGCA 1468
DB 1450 CAAGCTGCCAGTCAAGTGTGAGCGGCGCGAGGCGCTGTGAGAGAAAGAAATTTCCACTAA 1509
QY 1469 GTGAGATGTGTGAGTTTGGGTGCTGTCTGTGGAGGCTTCTCATATGACGAGGCTCC 1528
DB 1510 GTCTGAGTGTGAGTTTGGAAATCTTCTGTGGAAATTTACTCTTGTGGGAGAGTCC 1569
QY 1529 GTACCCATAAATGTCACTGAAAGAGGTGTGTGAGGCGGTGTGAGAGAGGCTATCCGCAATGGA 1588

QY 1529 GTACCTTAATGTCACCTGAAGAAGCTGTCCGAGCCCTGGAGAGGGGATCCCATGCA 1588
 Db 1570 TTATTCAGAAATTCCTCCCTGAAGGAGCTGTCCCTCGGGTGGAGAGAGGCTCAAGATGGA 1629
 QY 1589 ACCCCGAGGGGCTGTCCAGGCGCCGTCGACGTCTCTATGAGCGCTGTGGAGGAGCA 1648
 Db 1630 TGCCCCGAGCGGCTGCCCCGCGCCGACGTCTATGAAAGTCATGANAAGTGTGGACCTGGA 1689
 QY 1649 GCGCGCGCGCGCGGCGCCACCTTCGCGAAACTGGCGCGAGAGCTGG 1692
 Db 1690 CGCGCGCATGGCGCCCTCTCTCTACAGCTCCGAGAGAGCTTG 1733

RESULT 13

ABST73326 standard; DNA: 2442 BP.

ABST73326;

04-DEC-2002 (first entry)

DNA encoding human c-src isoform.

Chromosome aberration; oncogenic fusion protein; cancer; oncogene;
 proliferative disease; cellular protein isoform; heat shock protein 90;
 HSP-90; rheumatoid arthritis; cancer; hematopoietic disorder;
 T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.

Homo sapiens.

MO200269900-A2.

12-SEP-2002.

01-MAR-2002; 2002MO-US06518.

01-MAR-2001; 2001US-27251P.

(CONF-) CONFORMA THERAPEUTICS CORP.

Fritz LC, Burrows FJ;

WPI: 2002-698710/75.

P-PSDB; ABG95117.

Treating genetically-defined disease associated with chromosomal
 aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 diseases, involves administering an inhibitor of heat shock protein 90

Disclosure: Page 313-314; 389pp; English.

The invention describes a method of treating genetically-defined disease
 associated with chromosomal aberrations yielding oncogenic fusion
 proteins (I), treating cancerous cells containing (I) in a heterogeneous
 cell population, treating proliferative diseases associated with mutant
 protein or cellular protein isoforms (II) dependent on heat shock
 protein (HSP)-90, or selectively treating cells expressing (II)
 involving administering HSP90-inhibitor. The method is useful for
 treating genetically-defined disease with chromosomal aberration yielding
 oncogenic fusion protein, treating cancerous cells containing fusion
 protein in heterogeneous cell population, treating proliferative disease
 (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 p53), or selectively treating cells expressing mutant protein or cellular
 protein isoform in a patient heterozygous for (II). The method is useful
 for treating a disease e.g. haematopoietic disorder such as T or B cell
 lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML.

CC or a disease characterised by a solid tumour such as papillary thyroid
 carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 synovial sarcoma. The method is also useful for treating viral
 infections. This sequence encodes a human oncogenic protein.

SO Sequence 2442 BP: 474 A: 779 C: 708 G: 481 T: 0 other:

Query Match 26.4%; Score 527.8; DB 24; Length 2442;

Best Local Similarity 64.9%; Pred. No. 2.3e-96;

Matches 847; Conservative 0; Mismatches 447; Indels 11; Gaps 4;

QY 395 CTGGGCCCCGGGCGACCAAGTATATACCAATGAGACACACCGCCCAAGCCAGGGA 454
 Db 459 CTGGCATCCCGTACAAATGATATGCAATGATACCTTCACGCGACTGCCAGCAGA 518
 QY 455 GCTGGCTTCGCGAAGGCGAGCTGTACCATCTCTGAGGCGTGCAGAAACAGACTG 514
 Db 519 CTTGCCCTTCGCAAGAGACGCTGTCACCATTTGTGGCGTCCAGCAAGACCCCACTG 578
 QY 515 GTACCGGCTCAAGCACACACCACTGAGACAGAGGGGCTGTGACCTGGGCGCTGGC 574
 Db 579 GTACAAAGCCA---AAACAAAGTGGGCGCGTGAAGGCGATCATCCACGCAACTACGTC 635
 QY 575 GGAGCGGAGGCGCTCTCCGACACCCCAAGCTTACCTCATGCGCTG-CTTCCAGCGGA 633
 Db 636 GAACCGGGAGGGCGCTGAAGCGGGGTACCAACTACGCTCATGCGCTGAGTTCACGCGCA 695
 QY 634 AGATCTGGGGCGACAGAGGCTGTCCAGCAGCTGCGAGCTCCCGAGATGGGCTGTCTGG 693
 Db 696 AGATCAACAGGAGACAGGCTGTAGCGGCTTGTATACCGCGGAGACAGGCTGTCTGG 755
 QY 694 TGGCGAGTCCGCGCGCCACCGCGGCGACTACGTCTGTGCGAGCTTTGGCGCGACG 753
 Db 756 TGGCGAGAGACACCAACTACCGCGAGACTACAGCTGTGCGTGTGCGAGCGCAAGG 815
 QY 754 TCAATCCACTACCGGCTGTGACCGGCGACCGCCACCTCACAATGATGAGCGGTGTTCT 813
 Db 816 TGGAGCACTACCGGATATGATACATCCAGCAAGCTCAGCATGAGAGAGGTGTAAT 875
 QY 814 TCTGCAACCTCATGAGATGAGTGTGAGCATTTGACATTTGGAGGAGGCGGTATCTGACCA 873
 Db 876 TTGAGAACCTCATGACACTGTGTGAGGACATCACTACAGCAGAGATGAGTGTGATCGC 935
 QY 874 AGCTGTGAGACCAACCAAGCGGAACACGCGGACCAAGTGTGCGGAGAGAGCTGGCCAGG 933
 Db 936 GCTCATTTAAACCAAGGTATGAGGCGGCGACAGTGGCGGCGCATGATGATGCA 995
 QY 934 CGGGCTGTACTGAACTGACATGAGCATTTGATTTGGAGGAGCAGATGCGAGAGGAGT 993
 Db 996 GCGGCTGGCGCTGAACTGAAAGAGCTGAGCTGTGACAGACCATGCGAGAGGGAGT 1055
 QY 994 TTGAGCTGTCTGCAAGGCTGAGTACCTGGGCAAAAGGTGGCGGTGAAATATCAAGT 1053
 Db 1056 TCGGAGACGTGATGCTGTGGGATTTACGAGGAAACAAAGTGGCGCTCAATGATTAAGA 1115
 QY 1054 GTGATGTGACGCCAGGCGCTTCCTGAGCAGACGCGGCTCATGACAGATCAACACG 1113
 Db 1116 ACAGCGCACTGCGCAGGCGCTTCCTGCTGAACCTTCATGACAGCACTCGGCGATA 1175
 QY 1114 AGAACCTGTGCGTCTGCTGGGCTGATCTGTGACACG-----GGGCTTACATTTGTCA 1167
 Db 1176 GCAACCTGTGCGCTGCTGGGCTGATCTGTGAGGAGGAAGGCGGCGCTTACATGTCA 1235
 QY 1168 TGAGACAGTGACCAAGGAGGCACTGTGTAATTTTGGGAGCGCGGGTGCAGACCTCG 1227
 Db 1236 CTGAGTACATGCGCAAGGAGGAGGAGCTGTGTGACTACCTGCGGTTAGGGGCTGCACTGC 1295
 QY 1228 TGAACACCGCTGACGCTCTGCACTTTCTGTGACGTGGCGGAGGCGCATGAGTACTCG 1287
 Db 1296 TGCGGAGAGCTGTCTCTCAAGTCTCGTATGATGTCTGCGAGGCGCATGAGTACTCG 1355
 QY 1288 AGAGCAAGAAAGCTTGTGACCGGAGCTGGCGCGCGCAACATCTCTGCTCAGAGAGAC 1347

[illegible]

```

FT variation /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 6309, A)
FT /tag= a1
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 8753..8754, T)
FT /tag= a1
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 10797, T)
FT /tag= ak
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 10806, G)
FT /tag= al
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 11199, T)
FT /tag= am
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 11201, T)
FT /tag= an
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 11319, C)
FT /tag= ao
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 12238, C)
FT /tag= ap
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 14739, T)
FT /tag= aq
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT (replace 15375, C)
FT /tag= ar
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT WO200252018-A2.
FT 04-JUL-2002.
FT 19-DEC-2001; 2001WO-US48546.
FT 21-DEC-2000; 2000US-0741154.
FT (PEKE ) PE CORP NY.
FT Beasley EM, Shao W, Ketchum K, Di Francesco V;
FT WPI; 2002-583568/62.
FT P-PSDB; AAU99575.
FT New isolated human kinase proteins and genes, useful in developing
FT drugs, as well as for diagnosing, preventing or treating disorders
FT associated with defective cell signal transduction, e.g. cancer or
FT hematopoietic disorders
FT Claim 4; Fig 3A-F; 78pp; English.
XX
XX
XX CC The present invention relates to the isolation of a novel human kinase
XX related to the tyrosine kinase family, and polynucleotide sequences
XX encoding it. The gene encoding the kinase of the invention maps to
XX chromosome 19. The kinase may be used for identifying a modulator of
XX the kinase, an agent that binds to the kinase, or for identifying
XX other members of the family. The kinase may also be used to raise
XX antibodies which may be used in immunoassays or drug screening assays.
XX The sequences of the invention may be used for treating a disease or
XX condition mediated by a human kinase such as hematopoietic disorders
XX and other disorders associated with defective cell signal transduction
XX (e.g. cancer). The polynucleotide sequences encoding the kinase are
XX useful for isolating and purifying the kinase, and as probes or primers.
XX The present sequence encodes the human kinase of the invention.
XX
XX Sequence 16389 BP; 3511 A; 4901 C; 4515 G; 3462 T; 0 other;

```

Query Match 22.8%; Score 455; DB 24; Length 16389;
 Best Local Similarity 100.0%; Pred No. 11e-81;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1541 GTCACCTGAAGAAGCTGTCGAGGCGCTGAGAGAGGGGTACCGCATGAAACCCCGAGGG 1600
DB 13941 GTCACCTGAAGAAGCTGTCGAGGCGCTGAGAGAGGGGTACCGCATGAAACCCCGAGGG 14000
QY 1601 CTGTCCAGGCGCCGTCACGTCCTCATATGACAGCTGCTGGAGAGCCGCGCCGCG 1660
DB 14001 CTGTCCAGGCGCCGTCACGTCCTCATATGACAGCTGCTGGAGAGCCGCGCGCG 14060
QY 1661 GCCACCCCTCCCAAACTGGCCGAGAGAGCTGCGCGGAGCTACGAGTGCAGGTGCCCC 1720
DB 14061 GCCACCCCTCCCAAACTGGCCGAGAGAGCTGCGCGGAGCTACGAGTGCAGGTGCCCC 14120
QY 1721 ACCCTCCGTCACGAGGCGAGAGCCGAGGCTCCACCTTCGCCCGAAGCCAGAGCCCTG 1780
DB 14121 ACCCTCCGTCACGAGGCGAGAGCCGAGGCTCCACCTTCGCCCGAAGCCAGAGCCCTG 14180
QY 1781 ACCCCACCCGCTGGGGCCCTTGGCCGAGAGAGCCGAGAGTGTGAGAGTGGCGGTGG 1840
DB 14181 ACCCCACCCGCTGGGGCCCTTGGCCGAGAGAGCCGAGAGTGTGAGAGTGGCGGTGG 14240
QY 1841 GGCACCTGACCAAGGCCCAAGAGAGGTCACAGCGGCGCAATCTCTCTGTGTCACAGC 1900
DB 14241 GGCACCTGACCAAGGCCCAAGAGAGGTCACAGCGGCGCAATCTCTCTGTGTCACAGC 14300
QY 1901 AGGGCGTGGCCCTACGCTAGGGGCTCTGGGCGGCCGCTGACACCCAGACCTGCGAAGA 1960
DB 14301 AGGGCGTGGCCCTACGCTAGGGGCTCTGGGCGGCCGCTGACACCCAGACCTGCGAAGA 14360
QY 1961 TGATCGCCCGATTAAGACGATTTCTTAAGACTCTA 1995
DB 14361 TGATCGCCCGATTAAGACGATTTCTTAAGACTCTA 14395

```

```

RESULT 15
ABL11799
ID ABL11799 standard; cDNA; 2591 BP.
XX
XX AC ABL11799;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 29879.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX
XX OS pharmaceutical; gene; ss.
XX
XX XX Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX
XX PR 23-MAR-2000; 2000US-191637P.
XX
XX DR 11-JUL-2000; 2000US-0614150.
XX
XX PA (PEKE ) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PMD, Myers EW;
XX
XX PS WPI; 2001-656860/75.
XX
XX CC P-PSDB; ABB67696.
XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions
XX
XX Claim 1; SEQ ID NO 29879; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is

```


CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB16173-AB17022).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX

Sequence 2591 BP; 651 A; 733 C; 738 G; 469 T; 0 other;

Query Match 16.5%; Score 330.2; DB 23; Length 2591;
Best Local Similarity 57.5%; Pred. No. 8.2e-57;
Matches 637; Conservative 0; Mismatches 458; Indels 12; Gaps 2;

QY 597 GACCCAGCTCAGCCATCGCGTGTTCACAGGGAAGATCTCGGCGCCAGAGGCTGTC 656
DB 1413 GAGGTCAAACTGACGCGCATGCTTCATGCGCATTAACGGCGGATGAGCGCGAG 1472
QY 657 CACGAGCTGAGCGCTCCCGAGATGGGCTGTTCCTGGTGGGGAGTCCGGCGCACCC 716
DB 1473 CATCTGCTGACGCGCGCTGAGGATGTTTCTCTGCTGCGAGTCCACAAACTTCCC 1532
QY 717 GCGGACTACGCTCTGCTGAGCTTTGGCCGACGCTCATCCACTACCGCTGCTGCAC 776
DB 1533 GCGGACTACAGCGCTGCTGTTTCCATCCAGGTGAGCAGCTACCGGCTCAAGTAC 1592
QY 777 GCGGAGCGCCACTCACAATGATGAGCGCGTGTCTTCTGCAACTCATGAGCAGTGG 836
DB 1593 CTGAGAGAACAGCTGACATCGACGAGGAATACTTTGGAATTTGGCGCAGCTAGTG 1652
QY 837 GACATTATACAGCAAGGCGGCGTATCGACCAAGCTGGTGAGACCAAGCGGAA 896
DB 1653 GCGCACTACAGCGCGGCTGCGGATGAGCTGTCACCGCTGATTAATCTTCCAAAG 1712
QY 897 CACGGGACCAAGTCGCGC-----CGAGGAGAGCTGGCCAGGCGGCTGGTACTG 947
DB 1713 CTGGGCAAAACAGGTTTTCATTAACTCCAAAGATTTCTGGAGCAAGGCTGGTATC 1772
QY 948 AACCTGACGATTGACATTGGAGCAGACAGATCGAGAGGAGATTGGAGCTGTCTCG 1007
DB 1773 CCGTAACCGGAGCTGCGAGCTGCGCGAGAGCATCGCGCAAGGCGGATCGGTGATG 1832
QY 1008 CAGGCTGAGTACCTGGGGGCAAAAGTGGCGCTGAGAAATATCAAGTGTGATGACGCC 1067
DB 1833 CTGGGCTATTAACGACAAAGAGTGGCGCTCAAGATGCTGAAGACGAGGCGCGCTG 1892
QY 1068 CAGGCTTCTCTGAGCAGAGCGCGCTATGACGAAGATGCAACGAGCAAGAACTGTGCGT 1127
DB 1893 CAGAAGTTTCTGGCTGAGGCTCGGTGATGACACCGCTGAGCAGCAATTTGGTAAAG 1952
QY 1128 CTCTGGGCGTGAATCTGCA---CAAGGGCTGTACATTGTCAATGAGACGTTGAGCAAG 1184
DB 1953 TTTATGCGGCTTGTCTTACACAGCAAGCATCTATCTGTACGGAAATATATGAGTAAG 2012
QY 1185 GGCACCTGATGAACTTTCTGGGAGCCCGGGGTGAGCCCTCGTGAACAACCGCTCAGCTC 1244
DB 2013 GGATCACTGATTAATTAATCTGCGATGCGAGAGCAGACGACATTAACCAAAAAGATCAA 2072
QY 1245 CTGCAAGTTTCTCTGACAGTGGCGGAGGATGAGTACTGAGAGCAGAAAGCTTTGTG 1304
DB 2073 ATCATTTTTCCTACAGACACTGCTGCTGATGAGATATCTGGAAGCCAAAAGTTGTG 2132
QY 1305 CACCGGAGCTGGCGCGCCGCAACATCTGCTCAGAGGAGCTGTGGCCAAAGTCAAGC 1364
DB 2133 CATCGGATCTGGCAGCTGCGCAATGCTTATATCAGAGGATGCGTGAAGTATCG 2192
QY 1365 GACTTTGGCTGGGCAAAAGCGGAGGAGGCTAGACTCAAGCGCGCTGCCGTCAG 1424
DB 2193 GATTGGGTTTGGCAGCGGAGGAATGCTTACAAATCTTGAAGCTGGCAAGCTGCCATTA 2252
QY 1425 TGGACGGCGCCGAGGCTCTCAAAACAGGGAAGTTCACCAAGTCTGAGT 1484

DB 2253 TGGACGGCTCCGAGGACATTAAAAATGCGCTTTTCCAAACAATCAGCATGTGAGC 2312
QY 1485 TTTGGGCTGCTGCTGAGGAGTCTTCTCATATGAGAGGCGCTCCGTAACCTAAATGTCA 1544
DB 2313 TTTGGCTACTGCTATGAGGAAATCTACTCTTTGACGCGCTGCTATCCAGAAATCCA 2372
QY 1545 CTGAAGAGGTGTCGAGAGCGCGTGGAGAAAGGGGTACCGCATGGAACCCCGAGGCGCTG 1604
DB 2373 TTAGCCGATGTGTAAGACAGCTTGAAGTGGCTACAAATGGAAGCGCCGAGGCTGT 2432
QY 1605 CCAAGGCGCCGTGACGCTCTCATGAGCAGCTGCTGGAGGACAGACCCCGCGGCA 1664
DB 2433 CCAAGGAGATTTACGAATGATGATGCGGACAGCGCTGGAGCTTAATCCCGCAAGCGACC 2492
QY 1665 CCCTTCGCAAACTGCGCGGAGAAAGCTG 1691
DB 2493 ACATTCGCGGAGCTCAAGGTCAAACTG 2519

Search completed: August 1, 2003, 15:38:06
Job time : 676 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 15:09:19 ; Search time 5073 Seconds
(without alignments)
9581.895 Million cell updates/sec

Title: US-09-977-260-1
Perfect score: 2000
Sequence: 1 ctgcctcaagtgtgacg.....attctaagacttaaaaa 2000

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1071.8	53.6	2046	11 AK013606
2	926.2	46.3	1035	9 AL580639
3	924	46.2	1040	9 AL58929
4	909.6	45.5	1032	9 AL558805

C	5	867	43.4	906	10	BG744770	BG744770	602722891
C	6	857.2	42.9	1012	9	AL580543	AL580543	602722891
C	7	845.4	42.3	1043	12	B0054024	B0054024	602722891
C	8	830	41.5	949	9	AL567349	AL567349	602722891
C	9	811.4	40.6	1102	12	BM561134	BM561134	602722891
C	10	799	40.0	967	9	AL538511	AL538511	602722891
C	11	795.4	39.8	862	12	B1768288	B1768288	602722891
C	12	762	38.1	912	12	B1819535	B1819535	602722891
C	13	737.2	36.9	1014	12	BM925586	BM925586	602722891
C	14	707.6	35.4	733	13	B0617908	B0617908	602722891
C	15	707	35.4	986	13	BX369553	BX369553	602722891
C	16	693.4	34.7	842	12	B1910376	B1910376	602722891
C	17	693	34.6	712	12	BM681434	BM681434	602722891
C	18	691.4	34.6	1129	12	BM450510	BM450510	602722891
C	19	680.6	34.0	936	13	B1844285	B1844285	602722891
C	20	679.6	34.0	822	10	BG358801	BG358801	602722891
C	21	677.2	33.9	772	10	BG758878	BG758878	602722891
C	22	674.4	33.7	805	10	BG781641	BG781641	602722891
C	23	671.8	33.6	1119	12	BM925766	BM925766	602722891
C	24	661	33.1	708	12	BM715548	BM715548	602722891
C	25	660.4	33.0	702	12	BM674009	BM674009	602722891
C	26	641.2	32.1	732	12	B1769001	B1769001	602722891
C	27	640.4	32.0	837	12	B1912704	B1912704	602722891
C	28	626.4	31.3	651	13	B0582182	B0582182	602722891
C	29	624.6	31.2	741	12	BM719266	BM719266	602722891
C	30	621.4	31.1	729	9	AU135553	AU135553	602722891
C	31	621.2	31.1	767	12	B1911599	B1911599	602722891
C	32	608	30.4	953	12	B1488576	B1488576	602722891
C	33	602.4	30.1	605	13	B0070968	B0070968	602722891
C	34	602.2	30.1	729	9	AV702542	AV702542	602722891
C	35	599.6	30.0	931	12	B1549778	B1549778	602722891
C	36	598	29.9	735	10	BE779216	BE779216	602722891
C	37	597.6	29.9	847	10	BG396388	BG396388	602722891
C	38	595	29.8	607	10	BE257861	BE257861	602722891
C	39	593	29.6	593	13	B0073359	B0073359	602722891
C	40	592.2	29.6	631	10	BE254257	BE254257	602722891
C	41	589	29.4	613	10	BG396034	BG396034	602722891
C	42	587.6	29.4	937	13	B0438002	B0438002	602722891
C	43	577.8	28.9	1138	10	BE791829	BE791829	602722891
C	44	576.2	28.8	816	10	BE783096	BE783096	602722891
C	45	571	28.5	594	13	B0077460	B0077460	602722891

ALIGNMENTS

RESULT 1
AK013606
LOCUS
DEFINITION
Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone:2900029A15 product: megakaryocyte-associated
tyrosine kinase, full insert sequence.
ACCESSION
AK013606
VERSION
AK013606.1 GI:12851034
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalizing and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159


```
Db 508 CGTGCATCATCATACCTACCGTGTGTTTGCATGAGATGGGACCTTCACCTCATGANTGAGCC 567
Oy 807 GTGTTCTTCTGCAACCTCATGACATGGTGGAGCATTTACACCAAGAACAGGGCCCTATC 866
Db 568 GTGTTCTTCTTACCTGATGACATGGTGGAGCATTCACCAAGAACAGGGGCCATC 627
Oy 867 TGCACCAAGCTGTGTAGACCAAGAGGGAACACGGAGACCAAGTGGCGGAGGAGAGCTG 926
Db 628 TGCACCAAGCTGTGTAAAGCAAGAGGGAACAGGGGCCAAGTGTGCAGAGAGAGAGCTC 687
Oy 927 GCCAGGCGGGCGGTACTACCACTGCAGCATTTGACATTTGGAGACACAGATCGAGAG 986
Db 688 GCCAAGGCTGGCTGCTACTGACATCGACATCTGAGCATCTGGGAGACACAGATTTGAGAT 747
Oy 987 GGAGGTTTGGAGGCTGCTGACAGGAGTACCTGGGGGCAAAAGTGGCCGTGAGAT 1046
Db 748 GGGAGTTTGGAGCGCTCTACAGAGGTAGTACCTGGGACAGAAAGTGGCTGTGAGAT 807
Oy 1047 ATCAAGTGTATGTACAGCCAGGCTTCTTGCAGAGAGAGCGCCGTATGACGAAGATG 1106
Db 808 ATCAAGTGTATGTACAGCCAGGCTTCTTGCATGATGAGAGCGCTGATGACGAAGCTG 867
Oy 1107 CAACACGAACCTGTGTGCTGCTCTGGGCGTATCTGCACACAGGGGCTGTACTATTGTC 1166
Db 868 CAGCAGAGAACTGTGTGCTGCTCTGGGCGTATCTGCACACAGCGCTGTACTATTGTC 927
Oy 1167 ATGGAGCATGTGAGCAAGGGAACCTGTGTGAACCTTCTGGCGAGCCGGGGGTGAGCCCTC 1226
Db 928 ATGGAGCATGTGAGCAAGGGAACCTGTGTGAACCTTCTGGCGAGCCGGGGGTGAGCCCTC 987
Oy 1227 GTGAACACCGCTGACGCTCTGACAGTTTCTTCTGCACGTCGCGAGGCGATGAGTACTG 1286
Db 988 GTGAGCAGCTGTGACGCTCTGACAGTTTCTTCTGATGTCGTAAGGATGAGTAACTCTG 1047
Oy 1287 GAGAGCAAGAACTGTGTGACACCGGAGCTGGCGCGGCAACATCTGTGTCTGACAGAC 1346
Db 1048 GAGAGCAAGAACTGTGTGACACCGGAGCTGGCGCGGCAACATCTGTGTCTGACAGAC 1107
Oy 1347 CTGGGGCCCAAGGTCAGGAGCTTTGGGCTGGCCAAAGCGGAGCGGAGGGGCTAGACCA 1406
Db 1108 TTGGTGGCCCAAGGTCAGGAGCTTTGGGCTGGCCAAAGCGGAGCGGAGGGGCTGAGCTCA 1167
Oy 1407 AGCGGCTGCCCTGTCACAGTGTGACGCGCGCCGAGGCTCTCAAAACAGGGAGTTCACAGC 1466
Db 1168 AGCGGCTGCCAGTCAAGTGTGACGCGCGCACCTGAGGCTCTCAAAACGCGGCTTCTCCAGC 1227
Oy 1467 AAGTCGATGTCTGTGAGTTTGGGCTGCTCTGCGGAGGTCCTTCTCATATGAGAGGCT 1526
Db 1228 AAGTCGATGTCTGTGAGTTTGGGCTGCTCTGTTGGGAAGTCTTCTTATGGAAGAGCC 1287
Oy 1527 CCGTACCTTAATATGTCACTGAAGAGGTGTGCGAGGCGCGTGGAGAAAGGGTACCGCATG 1586
Db 1288 CCATACCCCAAGATGTGCTTAAGAGAGTTTCAAGAGGTGTGAGAAAGGTTTACCGCATG 1347
Oy 1587 GAACCCCGGAGGGGTGTCCAGGCGCCCTGTCACAGCTCTCATGAGAGAGTGTGAGAGCA 1646
Db 1348 GAGCGCGCGGATGGGTGCGCAGGCTGTGTGACACACCTCATGAGTGTGTGAGAGCA 1407
Oy 1647 GAGCGCGCGCGCGGCGCACCTTCCGCAACTGTGCGGAGAAAGTGTGCGGAGAGTACGC 1706
Db 1408 GAGCGCTGTGCGCGCGCACCTTCCGCAAAATATGAGAGTGTGCGGAGAGCTGAGTCCGC 1467
Oy 1707 AGTGCAGGTGCGCGAGCTGCGTCTGACAGGAGAGAGCGGAGGCTCCAGCTCGCGCCGA 1766
Db 1468 AGTGTGGTGTCTGTGCGCGCGCGCGTGGGAGACAGGAGGCTGAGGCTGAGTCCACACAG 1527
Oy 1767 AGCCAGGAGCCCTGAGCC 1785
Db 1528 AGCCAGGAGCCCTGATCC 1546
```

RESULT 2
AL580639/c

```
LOCUS AL580639 1035 bp mRNA linear EST 01-JUN-2003
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION AL580639
VERSION AL580639.2 GI:31318906
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1035)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 16, 2001 this sequence version replaced gi.12946853.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ007AF03NP1cluster-9238.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DJ007AF03NP1.
FEATURES
source
1..1035
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ007XR05"
/cell_line="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 166 a 327 c 304 g 213 t
ORIGIN
Query Match 46.3% Score 926.2; DB 9; Length 1035;
Best Local Similarity 97.3%; Fred. No. 5.9e-173;
Matches 960; Conservative 10; Mismatches 14; Indels 3; Gaps 3;
Oy 936 GCGTGTACTGAACCTGACATTTGATGGAGACAGATCGAGAGGAGATT 995
Db 1035 GCGTGTACTGAACCTGACATTTGATGGAGACAGATCGAGAGGAGATT 976
Oy 996 GGAGCTGTCTGCAAGGAGTACTGTGGGCAAAAGTGGCGCTGAAGATATCAAGTGT 1055
Db 975 GGAGCTGTCTGCAAGGAGTACTGTGGGCAAAAGTGGCGCTGAAGATATCAAGTGT 916
Oy 1056 GATGTGAGAGCCGAGGCTTCTGTGAGACAGAGCGCGCTCATGAGCAAGATGCAACACAG 1115
Db 915 GATGTGAGAGCCGAGGCTTCTGTGASAGAGCGCGCTCATGAGCAAGATGCAACACAG 856
Oy 1116 AACCTGGGCTGTCTGCGGAGTATCTGTGACCAAGAGGCTGTACATTTGATGAGAGAC 1175
Db 855 AACCTGGGCTGTCTGCGGAGTATCTGTGACCAAGAGGCTGTACATTTGATGAGAGAC 796
Oy 1176 GTGAGCAAGGCGCAACTGTGTGACTTCTGTGCGGAGCCGGGTCGAGCCCTGTGAACAC 1235
Db 795 GTGAGCAAGGCGCAACTGTGTGACTTCTGTGCGGAGCCGGGTCGAGCCCTGTGAACAC 736
Oy 1236 GCTCAGCTCTGCAATTTCTCTGACAGTGGCCGAGGAGCATGAGTACTGTGAGAGCAAG 1295
Db 735 GCTCAGCTCTGCAATTTCTCTGACAGTGGCCGAGGAGCATGAGTACTGTGAGAGCAAG 676
Oy 1296 AAGCTTGTGCAACCGGAGCTGGCGCGCGCAACATCTGTGTCTGAGAGAGACTGTGGCC 1355
```

```

Db      675 AAGCTGTGCACGCGACCTGGCCGCCGCGCAACATCCTGTTCTCAGAGAACTGTGGCC 616
QY      1356 AAGGTGACGACTTTGGCCCTGCGCAAAAGCGAGAGAGGGGCTAGACTCAAGCCGGCTG 1415
Db      615 AAGTTCAGGACACTTTGGCCCTGGCCAAAGCCAGAGAGGGGCTAGACTCAAGCCGGCTG 556
QY      1416 CCGCTCAAGTGAAGCGGCGCCGAGGCTCTCAACACGGGAAGTTCAACCAAGTGGAT 1475
Db      555 CCGCTCAAGTGAAGCGGCGCCGAGGCTCTCAACACGGGAAGTTCAACCAAGTGGAT 496
QY      1476 GTCTGAGTTTGGGGGCTGCTGAGAGTCTTCTCATATGAGAGGCTCCGTTACCT 1535
Db      495 GTCTGAGTTTGGGGGCTGCTGAGAGTCTTCTCATATGAGAGGCTCCGTTACCT 436
QY      1536 AAATGTCACTGAAAGAGTGTGAGAGGCGGTGAGAGAGGGGTACCGCATGAAACCCCG 1595
Db      435 AAATGTCACTGAAAGAGTGTGAGAGGCTTGAAGAGGGGTACCGCATGAAACCCCG 376
QY      1596 GAGGCTGTCCAGGCGCCGCTGACGCTCTCATAGAGCAGTGTGGAGAGGAGCCCGCC 1655
Db      375 GAGGCTGTCCAGGCGCCGCTGACGCTCTCATAGAGCAGTGTGGAGAGGAGCCCGCC 316
QY      1656 CGCGGCGCAACCTTCGCGAAACGTGGCGGAGAACTGGCCCGGAGCTAGCAGTGCAGGT 1715
Db      315 CCGCGGCGCAACCTTCGCAAACTGGCGGAGAGAGTGGCCCGGGA-CTACGCACTGAGGT 257
QY      1716 GCCCGAAGCTTCGCTTCAGAGGAGAGAGCGGAGCTGCACCTCGCCCGCAAGCCAGAG 1775
Db      256 GCCCGAAGCTTCGCTTCAGAGGAGAGAGCGGAGCTGCACCTCGCCCGCAAGCCAGAG 197
QY      1776 CCTTGAACCCAGCCCGGTGGGGGCTTGGCCCAAGAGAGCCAGAGAGTGAAGTGGGCG 1835
Db      196 CCTTGAACCCAGCCCGGT-GGGCGCTTGGCCCGCAAGAGAGCCAGAGAGTGAAGTGGGCG 138
QY      1836 GTGGGGGCACTGACAGGCGCCAGAGAGGGGTCCAGGGGGGCAAGTCAATCCCTCGGTGGCC 1895
Db      137 GTGGGGGCACTGACAGGCGCCAGAGAGGGTCCAGGGGCAAGTCAATCCCTCGGTGGCC 78
QY      1896 ACAGAGGGGCTGGCCACGCTAGGGG 1922
Db      77 ACAGCA-GGGCTGGCCACGCTAGGGG 52

RESULT 3
AL558929 1040 bp mRNA linear EST 31-MAY-2003
LOCUS AL558929 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ007YK05 5-PRIME, mRNA sequence.
ACCESSION AL558929
VERSION AL558929.2 GI:31283062
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1040)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12903930.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by life technologies, a
division of invitrogen. This sequence belongs to sequence cluster
9238.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ007AF03QPl&cluster=9238.r. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ007AF03QPl.

```

```

FEATURES
    source          Location/Qualifiers
                    1..1040
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="CS0DJ007YK05"
                    /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                    /clone_1ib="JURKAT"
                    /note="1st strand cDNA was primed with a NotI-oligo(dT)
                    primer. Five prime end enriched, double-strand cDNA was
                    digested with Not I and cloned into the Not I and EcoR V
                    sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT      188 a      341 c      338 g      162 t      11 others
ORIGIN
Query Match      46.2% Score 924; DB 9; Length 1040;
Best Local Similarity 99.2% Pred. No.1.6e-172;
Matches 958; Conservative 2; Mismatches 2; Indels 4; Gaps 3;

1 CTCGCTCCAAAGTTGTGACAGCGGAGCGGCTCGGGGAGTGCACGCGGCTCGGAGGCC 60
Db      77 CTGCTCCAAAGTTGTGACAGCGGAGCGGCTCGGGGAGTGCACGCGGCTCGGAGGCC 136
QY      61 TCTTGGGGGCGGGCGCGGG--GGGCTCGGGGCGGCCCTGTAGCAGAAAAAGAGAA 118
Db      137 TCTTGGGGGCGGGCGGGCGGGGCTCGGGGCGGCCCTGTAGCAGAAAAAGAGAA 196
QY      119 CAGAGCTGGTCCAGTGGGACCCAGAGTCTCTACTCTGTGCGACGCGCTGGCTGTGG 178
Db      197 CAGAGCTGGTCCAGTGGGACCCAGAGTCTCTACTCTGTGCGACGCGCTGGCTGTGG 256
QY      179 CAGGCAATTCACAGCTCCCGCACTGTGACCACTCTCAGTGTGCTCAGCTCAGCTC 238
Db      257 CAGGCAATTCACAGCTCCCGCACTGTGACCACTCTCAGTGTGCTCAGCTCAGCTC 316
QY      239 AGTTTCCCTCTGGGGGCGATGGCGGGGCGAGGCTCTGTGTTCTGTGGGCGAATTTCA 298
Db      317 AGTTTCCCTCT-GGGGGCGATGGCGGGGCGAGGCTCTGTGTTCTGTGGGCGAATTTCA 375
QY      299 CGGCTGTGATTCGTGAGGAAGTTCGCGGGGTGAGACCCCGGTTCTCCGAGACCTGGCA 358
Db      376 CGGCTGTGATTCGTGAGGAAGTTCGCGGGGTGAGACCCCGGTTCTCCGAGACCTGGCA 435
QY      359 -CCGCCCTCCGCTCAGCCAGATGCCAAGAGAGGCTGGGCCCGCGGACCCAGTGTAT 418
Db      436 CCGCCCTCCGCTCAGCCAGATGCCAAGAGAGGCTGGGCCCGCGGACCCAGTGTAT 495
QY      419 CACCAATGCGAGACACCGCGCCCAAGCCAGGGGAGCTGGCTTCGCAAGGGGAGCT 478
Db      496 CACCAATGCGAGACACCGCGCCCAAGCCAGGGGAGCTGGCTTCGCAAGGGGAGCT 555
QY      479 GGTACACATTCCTGAGAGGCTGGGAGAAAGAAAGACTGTACCGGTCAACACACACG 538
Db      556 GGTACACATTCCTGAGAGGCTGGGAGAAAGAAAGACTGTACCGGTCAACACACACG 615
QY      539 TGGACAGAGAGGCTCTGAGAGTGGGCGCTGCGGAGCGGAGAGGAGCCCTTCCGACA 598
Db      616 TGGACAGAGAGGCTCTGAGAGTGGGCGCTGCGGAGCGGAGAGGAGGAGCCCTCCGACA 675
QY      599 CCGCAAGCTCAGCTCATGCGCGTGTTCACAGGGGAAGATCTGGGCGCAGAGGCTGTCA 658
Db      676 CCGCAAGCTCAGCTCATGCGCGTGTTCACAGGGGAAGATCTGGGCGCAGAGGCTGTCA 735
QY      659 GCAAGTGCAGGCTCCGAGAGATGAGGCTGTTCGAGGCGGAGAGTCCGCGCACCGCGG 718
Db      736 GCAAGTGCAGGCTCCGAGAGATGAGGCTGTTCGAGGCGGAGAGTCCGCGCACCGCGG 795
QY      719 CGACTACGCTCTGTGCTGAGCTTTGGCGGAGCTCATCACTACCGGCTGTGACCG 778
Db      796 CGACTACGCTCTGTGCTGAG-CTTTGGCGGAGCTCATCACTACCGGCTGTGACCG 854

```

QY 779 CGACGCGACCTCACAAATCGATGAGCGCGTGTCTTCTGCAACCTCATGAGCATGTGGA 838
|||||
Db 855 CGAGCGCACCTCACAAATCGATGAGCGCGTGTCTTCTGCAACCTCATGAGCATGTGGA 914
QY 839 GCATTACAGCAAGGACGAGCGCGTATCTGCAACCAAGCTGGTGAACCAAGCGGAACA 898
|||||
Db 915 GCATTACAGCAAGGACGAGCGCGTATCTGCAACCAAGCTGGTGAACCAAGCGGAACA 974
QY 899 CGGAGCAAGTGGCGGAGGAGGAGCTGGCGGAGGCGGCGTGGTACTGAACTGACGCA 958
|||||
Db 975 CGGAGCAAGTGGCGGAGGAGGAGCTGGCGGAGGCGGCGTGGTACTGAACTGACGCA 1034
QY 959 TTTGAC 964
|||||
Db 1035 TTTGAC 1040

RESULT 4
AL558805 1032 bp mRNA linear EST 31-MAY-2003
LOCUS AL558805 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens CDNA clone CS0DJ015Y001 5-PRIME, mRNA sequence.
ACCESSION AL558805
VERSION AL558805.2 GI:31280603
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1032)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12903685.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9238.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015AH01P1&cluster=9238.r. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ015AH01P1.
FEATURES
source location/Qualifiers
1..1032
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015Y001"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_11b="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-poly(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 187 a 338 c 334 g 164 t
ORIGIN

Query Match 45.5% Score 909.6 DB 9: Length 1032;
Best Local Similarity 99.1% Pred. No. 1.1e-169;
Matches 943; Conservative 3; Mismatches 2; Indels 4; Gaps 3;

QY 1 CTCGCTCCAAAGTTGTGACGCGCGGAGCGGCTCGGAGTGTGACCGCGCTCGCGAGGCC 60
|||||
Db 80 CTCGCTCCAAAGTTGTGACGCGCGGAGCGGCTCGGAGTGTGACCGCGCTCGCGAGGCC 139
|||||
QY 61 TCTCTGGGGGGGGGGGGGG--GCGGCTCGGGGGGGCCCCCTGAGCAAGAAACAGAGAACA 118
|||||

Db 140 TCTCTGGGGGGGGGGGGGGCGGCGGCTCGGGGGGCGCCCTGAGCAAGAAACAGAGAACA 199
QY 119 CGAGCTCGGCTCAGTGGACACCCAGCTCCCTACCTCCTGCTGCGACCGCGCTGGG 178
|||||
Db 200 CGAGCTCGGCTCAGTGGACACCCAGCTCCCTACCTCCTGCTGCGACCGCGCTGGG 259
QY 179 CAGGCCATTTCCAGGCTCCCGACCTGTCAGCACTTGTCTCAGTGTCTCTCAGTGTCT 238
|||||
Db 260 CAGGCCATTTCCAGGCTCCCGACCTGTCAGCACTTGTCTCAGTGTCTCTCAGTGTCT 319
QY 239 AGTTTCCTCTGCGGGGGGCGATGCGGGGCGAGGCTCTCTGTTTCTTGCGGGGCAATTTCA 298
|||||
Db 320 AGTTTCCTCTCT--GGGGGCGATGCGGGGCGAGGCTCTCTGTTTCTTGCGGGGCAATTTCA 378
QY 299 CGGCTGTGATTTCTGCTGAGGAATTTCCCGGGGTGAGCCCGCTTCTTCCGAGCTTGCA 358
|||||
Db 379 CGGCTGTGATTTCTGCTGAGGAATTTCCCGGGGTGAGCCCGCTTCTTCCGAGCTTGCA 438
QY 359 CCCCTCCCGCTCTCAGCGAGGATGCCAAGAGGCGCTGGGCGCCGCGGACCCAGTGTAT 418
|||||
Db 439 CCCCTCCCGCTCTCAGCGAGGATGCCAAGAGGCGCTGGGCGCCGCGGACCCAGTGTAT 498
QY 419 CACCAATGCGAGCACACCCGCGCCCAAGCGAGGAGCTGGCTTCCGCAAGGCGAGCT 478
|||||
Db 459 CACCAATGCGAGCACACCCGCGCCCAAGCGAGGAGCTGGCTTCCGCAAGGCGAGCT 558
QY 479 GGTACCAATCTCTGAGGCGCTGGGAGAAAGAGCTGTACCCGCTCAAGCAACCAACAG 538
|||||
Db 559 GGTACCAATCTCTGAGGCGCTGGGAGAAAGAGCTGTACCCGCTCAAGCAACCAACAG 618
QY 539 TGGACAGAGGAGGCGTGTGAGCTGGGCGCTGGGAGCGGAGCGGAGGCGCTTCCGCA 598
|||||
Db 619 TGGACAGAGGAGGCGTGTGAGCTGGGCGCTGGGAGCGGAGCGGAGGCGCTTCCGCA 678
QY 599 CCCCAAGCTCAGCTCTCATGCTGCTGTTCACGCGGAAGATCTGGGCGAGAGGCTGTCCA 658
|||||
Db 679 CCCCAAGCTCAGCTCTCATGCTGCTGTTCACGCGGAAGATCTGGGCGAGAGGCTGTCCA 738
QY 659 GGAGTGGAGGCTCCCGAGGATGGGCTGTCTGTTGGGGGAGTCCGGGCGCACCGCG 718
|||||
Db 738 GGAGTGGAGGCTCCCGAGGATGGGCTGTCTGTTGGGGGAGTCCGGGCGCACCGCG 798
QY 719 CGACTACGCTCTGCTGCTGAGCTTGTGGCGCGAGCTCATCTACCTACCGGCTCTCACCG 778
|||||
Db 799 CGACTACGCTCTGCTGCTGAGCTTGTGGCGCGAGCTCATCTACCTACCGGCTCTCACCG 858
QY 779 CGAGCGCACCTCACAAATCGATGAGCGCGTGTCTTCTGCAACCTCATGAGCATGTGGA 838
|||||
Db 859 CGAGCGCACCTCACAAATCGATGAGCGCGTGTCTTCTGCAACCTCATGAGCATGTGGA 918
QY 839 GCATTACAGCAAGGACGAGGCGGCTATCGCAACGCTGGTGAACCAAGCGGAACA 898
|||||
Db 919 GCATTACAGCAAGGACGAGGCGGCTATCGCAACGCTGGTGAACCAAGCGGAACA 978
QY 899 CGGAGCAAGTGGCGGAGGAGGAGCTGGCGGCGGCGTGGTACTGTAAC 950
|||||
Db 979 CGGAGCAAGTGGCGGAGGAGGAGCTGGCGGCGGCGTGGTACTGTAAC 1029

RESULT 5
Bg744770/c 906 bp mRNA linear EST 15-MAY-2001
LOCUS Bg744770/c 602722891n1 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:4849175 3',
DEFINITION mRNA sequence.
ACCESSION Bg744770
VERSION Bg744770.1 GI:14055423
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 906)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.

QY	679	ATGGGCTGTCTCTGGTGGGGAGTCCGGCCGACCCGGGCGACTACTGCTCTGTGCTGTA	738
QY	679	ATGGGCTGTCTCTGGTGGGGAGTCCGGCCGACCCGGGCGACTACTGCTCTGTGCTGTA	738
Db	301	ATGGGCTGTCTCTGGTGGGGAGTCCGGCCGACCCGGGCGACTACTGCTCTGTGCTGTA	360
QY	739	GCTTTGGGCGGACGCTCATCCACTACCGGTCCTCACCGGACGGGACCCCTCACATTCG	798
Db	361	GCTTTGGGCGGACGCTCATCCACTACCGGTCCTCACCGGACGGGACCCCTCACATTCG	420
QY	799	ATGAGGCGGTGTCTTCTTGCAACTCATATGACATGTGTGGAGCATTTACAGCAAGCAAG	858
Db	421	ATGAGGCGGTGTCTTCTTGCAACTCATATGACATGTGTGGAGCATTTACAGCAAGCAAG	480
QY	859	GCGCATCTGCACCAAGCTGGGTGAACCAAGGCGGAAACAGGGGACCAAGTGGCGGAG	918
Db	481	GCGCATCTGCACCAAGCTGGGTGAACCAAGGCGGAAACAGGGGACCAAGTGGCGGAG	540
QY	919	AGGAGCTGGCCAGGGCGGGCTGTTACTGAACTTCAGCTCAGCATTTTGAATTGGAGCACA	978
Db	541	AGGAGCTGGCCAGGGCGGGCTGTTACTGAACTTCAGCTCAGCATTTTGAATTGGAGCACA	600
QY	979	TCGGAGAGGAGAGACTTTGGAGCTGTCTTCGACAGGCTGAATACCTGGGGCAAAAGTGGCCG	1033
Db	601	TCGGAGAGGAGAGACTTTGGAGCTGTCTTCGACAGGCTGAATACCTGGGGCAAAAGTGGCCG	660
QY	1039	TGAAGAAATCAAGTGTATGTATGACAGCCGAGCCCTTCCTGGAGCAGAGCGGCGTCATGA	1098
Db	661	TGAAGAAATCAAGTGTATGTATGACAGCCGAGCCCTTCCTGGAGCAGAGCGGCGTCATGA	720
QY	1099	CGAAGATCAACACAGAG- AACCTGGTGGCTCTCTGGGCGGTATCCTCACAGAGGCGCTG	1157
Db	721	CGAAGATCAACACAGAGAACCTGGTGGCTCTCTGGGCGGTATCCTCACAGAGGCGCTG	780
QY	1158	TACATGTTCATGAGGACAGTGAAGCAAGGCAAACTT- GGTGAACCTTTCTGGGACCCGGG	1218
Db	781	TACATGTTCATGAGGACAGTGAAGCAAGGCAAACTT- GGTGAACCTTTCTGGGACCCGGG	840
QY	1217	--TCGAGCCCTCGTGACACCGGCTCAGCTCCTGCAAGTTTCTCTGACAGTGGCC	1268
Db	841	GTGAGGCCCTCGTGAAACACCGGCTCANTCTCTGCAAGTTTCTCTGACAGTGGCC	894

RESULT 8
AL567349/c

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL567349 949 bp mRNA linear EST 31-MAY-2003
AL567349 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CSDDF027Y13 3-PRIME, mRNA sequence.
AL567349
AL567349.2 GI:31290254
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 949)
L1.W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12520618.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/cluster.cgi?seq=CSDDF027CB07NP1&cluster=9238.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSDDF027CB07NP1.
Location/Qualifiers
I. 949

FEATURES
SOURCE

[illegible]

Db	170	GGGSCCTTGGGCCCCCAGAGACCCAGAGACAGTGTGAGATGTGGCGGCTGTGGGGCACTGTGACAG	112
Qy	1853	GCCCAAGGAGGGGTCCAGGCGGGGCAAGTCATCTCTGTGTGCCACAGCAGGGGCTGGCCC	1912
Db	111	GGCCAAAGAGGGCCAGCGAGCGAGATCAACMCCT-GTGGCCACMGAGGGGTGNCCCA	53
Qy	1913	ACGTAGGGGGCTCTTGGGGGCGCCGTGTGACACCCAGACCC	1951
Db	52	CGTAGGGGNCCTTGGGGCGBCGCTGTGACACCCAGAC	14
RESULT 9			
LOCUS	BM561134	1102 bp	mRNA linear EST 20-FEB-2002
DEFINITION	AGENCOURT_6566231 NIH_MGC_67 Homo sapiens	cdna clone	IMAGE:5550935
ACCESSION	BM561134		
VERSION	BM561134.1	GI:18806135	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://limage.lnl.gov Plate: LNLMI2263 row: n column: 24 High quality sequence start: 32 High quality sequence stop: 611. Location/Qualifiers		
FEATURES			
source	1. 1102		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5550935"		
	/tissue_type="retinoblastoma"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NIH_MGC_67"		
	/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NCI; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies"		
BASE COUNT	199 a 377 c 354 g 172 t		
ORIGIN			
Query Match	40.6%; Score 811.4; DB 12; Length 1102;		
Best Local Similarity	95.6%; Pred. No. 3e-150;		
Matches	856; Conservative 0; Mismatches 36; Indels 3; Gaps 2;		
Qy	18	AGCCGGGACCGCCTCGGGGTGTGACAGCCGCTGCGAGAGCCCTCTGTGGGGCGGGCGCG	77
Db	44	AGCCGGGACCGCCTCTCGGGGTGTGACAGCCGCTGCGAGAGCCCTCTGTGGGGCGGGCGCG	103
Qy	78	GG--GGGGCTGGGGGGGGCGCCCGGAGAGAGAAACAGAGAACAGCAGGCTGGTCCAGTG	135
Db	104	GGGGGGGGCTGGGGGGGGCGCCCGCTGAGAGAGAAACAGAGAACAGCAGGCTGGTCCAGTG	163
Qy	136	GCACCCAGCTCTCTACTCTCTGTGTGACAGCCGCTGAGGCTGTGGCAGGCAATTCACACGT	195
Db	164	GCACCCAGCTCTCTACTCTCTGTGTGACAGCCGCTGAGGCTGTGGCAGGCAATTCACACGT	223
Qy	196	CCCCGACTGTACCACTTGTCTCAGTGTGCTCTTCACCTGCTCTCAATTCCTCTGAGGGGG	255

Db	224	CCCGAGCTGTGACAACTTGGTCGTAAGTGTCCTCTCCTCACTGCCTCAATTTGCCCT-66666	282
OY	256	CGATGGCGGGGCGAGAGCTCTCTGGATTGCTGTGGGCGCATTTTCACGGCTGTGATTCTGCTG	315
Db	283	CGATGGCGGGGCTGTAGGCTCTGTGTTCTGTGGCGGCAATTTCACGGCTGTGATTCTGCTG	342
OY	316	AGGAACCTTCCCCGGGTGAGCCCCGCTTCTCCGAGCCTGGACACCCTCCGTTCTAG	375
Db	343	AGGAACCTTCCCCGGGTGAGCCCCGCTTCTCCGAGCCTGGACACCCTCCGTTCTAG	402
OY	376	CCAGGATGCCAACAAGAGGGGCTGTGGGCCCCGGGACCCCACTGTATATCACCAAATGCCAGACA	435
Db	403	CCAGGATGCCAACAAGAGGGGCTGTGGGCCCCGGGACCCCACTGTATATCACCAAATGCCAGACA	462
OY	436	CCCGGCCCCAAGCAGAGGGAGGCTGTGGCTTCCGCAAGAAGGGAGAGTGATACACTTCCTGGAGG	495
Db	463	CCCGGCCCCAAGCAGAGGGAGGCTGTGGCTTCCGCAAGAAGGGAGAGTGATACACTTCCTGGAGG	522
OY	496	CCTGGAGAACAGAAGCTGTGTACCGCGCTCAAGCACACACCAAGTAGAGGAGGAGGCTGC	555
Db	523	CCTGGAGAACAGAAGCTGTGTACCGCGCTCAAGCACACACCAAGTAGAGGAGGAGGCTGC	582
OY	556	TGGCAGCTGTGGGGCGCTGTGGGGAGGCGGGAGGCGCTTCCGAGACCCCAGCTCAACCTCA	615
Db	583	TGGCAGCTGTGGGGCGCTGTGGGGAGGCGGGAGGCGCTTCCGAGACCCCAGCTCAACCTCA	642
OY	616	TGCCCTGGTTCCACAGGGAAGATCTGGGGCCAGAGAGGCTGTCCAGAGCTGAGGCTCCCG	675
Db	643	TGCCCTGGTTCCACAGGGAAGATCTGGGGCCAGAGAGGCTGTCCAGAGCTGAGGCTCCCG	702
OY	676	AGGATGGGCTGTCTCTGGTGGCGGAGTCCGGCGCACCCGCGGACTACCTCTGTGCG	735
Db	703	AGGATGGGCTGTCTCTGGTGGCGGAGTCCGGCGCACCCGCGGACTACCTCTGTGCG	762
OY	736	TGAGCTTTGGCCCGCAGCTATCCACTACCGGTGCTCACCOCGAGCGGACCACTCAACA	795
Db	763	TGAGCTTTGGCCCGCAGCTATCCACTACCGGTGCTCACCOCGAGCGGACCACTCAACA	822
OY	796	TGAGTGAAGGCGGTGTTCTTGCAACCCATGAGACATGTGGAGCATTAACGAAGAGACA	855
Db	823	TGAGTGAAGGCGGTGTTCTTGCTGCACTCCCGGACATGTGGAGCATTAACGAAGAGAAA	882
OY	856	AGGGCGCTATCTGCACCAAGCTGTGTGAGACCAAGGCGGAACACAGCGACCAAGTIC	910
Db	883	GCGGCGCTATCTGCACCAAGCTGTGTGAGACCAAGGCGGAACACAGCGACCAAGTIC	937
RESULT 10			
AL538511		967 bp mRNA linear EST 31-MAY-2003	
LOCUS	AL538511	Homo sapiens FETAL BRAIN Homo sapiens cdna clone	
DEFINITION	CSDDF027YI13.5-PRIME, mRNA sequence.		
ACCESSION	AL538511		
VERSION	AL538511.2	GI:31263101	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 967)		
JOURNAL	Li,W.B., Gruber,C., Jesse,J., and Polays,D.		
COMMENT	Full-length cDNA libraries and normalization		
	Unpublished		
	On Feb 13, 2001 this sequence version replaced gi:12802004.		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	Llibrary was constructed by Life Technologies, a division of		
	Invitrogen. This sequence belongs to sequence cluster 9238.r For		
	more information about this cluster, see		
	http://www.genoscope.cns.fr/		
	cgi-bin/cluster.cgi?seq=CSDDF027CB07Qp1&cluster=9238.r. Contact :		

Feng Liang Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com/>
 Faraday Avenue Genoscope sequence ID: CS0DF027CB070P1.

FEATURES

Location/Qualifiers

1..967

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DF027YD13"

/tissue_type="FETAL BRAIN"

/dev_stage="fetal"

/note="Organ: brain; Vector: pCMVSPORT-6; 1st strand cDNA was primed with a NotI-oligo(5') primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 193 a 297 c 330 g 147 t

Query Match 40.0%; Score 799; DB 9; Length 967;
 Best Local Similarity 97.0%; Pred. No. 8.1e-148;
 Matches 814; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

287 GCGGGCATTTACAGGCTGTGATTTCTGCTGAGGAATCTCCCGGGTGAGCCCGCTTCT 346
 129 GGGGGCGGGCGGGGGCGGCTCTGGGGGGCGCCCTGAGAGGTGAGCCCGCTTCT 188
 347 CCGAGCTGGACACCCCTCCCTGCTGCTGAGGAGATGCAACAGAGCGCTGGGCGCGG 406
 189 CCGAGCTGGACACCCCTCCCTGCTGCTGAGGAGATGCAACAGAGCGCTGGGCGCGG 248
 407 CACCAATGTATACCAATGCGACACACCCCGCCCAAGCCAGGAGAGCTGGCTTCCG 466
 249 CACCAATGTATACCAATGCGACACACCCCGCCCAAGCCAGGAGAGCTGGCTTCCG 308
 467 CAAGGGAGAGTGTGATCCTGAGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 526
 309 CAAGGGAGAGTGTGATCCTGAGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 368
 527 GCACACACAGTGTGAG 586
 369 GCACACACAGTGTGAG 428
 587 CTTCTCGGAG 646
 429 CTTCTCGGAG 488
 647 GAGAGCTGTGAG 706
 489 GAGAGCTGTGAG 548
 707 GCGCCACCCCGGAG 766
 549 GCGCCACCCCGGAG 608
 767 CGTGTGTGAG 826
 609 CGTGTGTGAG 668
 827 GAGAGTGTGAG 886
 669 GAGAGTGTGAG 728
 887 AAAGCGGAAACAGCGGAG 946
 729 AAAGCGGAAACAGCGGAG 788
 947 GAACTGTGAG 1006
 789 GAACTGTGAG 848
 1007 GCAAGGTGAG 1066

Db 849 GCAGGGTGTGAG 908
 Oy 1067 CCAGGCTTCTGAG 1125
 Db 909 CCAGGCTTCTGAG 967

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

EMAIL

ISSUE

CDNA

LIBRARY

ARRAYED

BY

THE

I.M.A.G.E.

CONSORTIUM

(LNL)

DNA

SEQUENCING

BY

INCYTE

GENOMICS

INC.

CLONE

DISTRIBUTION

MGC

CLONE

DISTRIBUTION

INFORMATION

CAN BE

FOUND

THROUGH

THE

I.M.A.G.E.

CONSORTIUM

LNL

AT

PLATE

LIBRARY

1516

ROW

C

COLUMN

11..862
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5205831"
 /lab_host="DH10B"
 /clone_lib="NIH-MGC_122"
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; site: 1; NotI; Site: 2; EcoRV (destroyed); RNA source anonymous pool of 24-week female lung, 16-week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC library."

BASE COUNT 143 a 298 c 281 g 140 t

Query Match 39.8%; Score 795.4; DB 12; Length 862;
 Best Local Similarity 98.1%; Pred. No. 4e-147;
 Matches 847; Conservative 0; Mismatches 11; Indels 5; Gaps 4;

30 CTGCGGCTGTGACCGGCTGCGGAGAGCCCTCTGGGGCGGCGCGG--CGGGCTCG 87
 1 CTGCGGCTGTGACCGGCTGCGGAGAGCCCTCTGGGGCGGCGGCGGCTCG 60
 88 GGGGGCGGCGGCTGTGACCGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 147
 61 GGGGGCGGCGGCTGTGACCGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 148 CTACCTCTGTGACCGGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 207
 121 CTACCTCTGTGACCGGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 208 CCACTTGTCTGAGTGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 267

|||||
Db 181 CCACTTGTCTAGTGTGCTCTCACTGCTCACTGTTCTCTCT -GGGGGCGATGCGGGG 239
QY 268 GAGGCTCTGTTCTCTGCGGGGCAATTTACAGGCTGTGATCTGCTGAGGAATCTCCC 327
Db 240 GAGGCTCTGTTCTCTGCGGGGCAATTTACAGGCTGTGATCTGCTGAGGAATCTCCC 299
QY 328 GGGGAGACCCCGCTCTCTCGAGAGCTGGACACCCCGCTCTCGAGAGAGATGCCAA 387
Db 300 GGGGAGACCCCGCTCTCTCGAGAGCTGGACACCCCGCTCTCGAGAGATGCCAA 359
QY 388 CGAGGCTGTGGGCGCGGGGACCCAGTGTATCACCAATGAGAGACACCGCGCCCAAGC 447
Db 360 CGAGGCTGTGGGCGCGGGGACCCAGTGTATCACCAATGAGAGACACCGCGCCCAAGC 419
QY 448 CAGGGAGCTGGCTTCCGCAAGGGGCACTGTGACATCTCTGAGGCTTCCGAGACA 507
Db 420 CAGGGAGCTGGCTTCCGCAAGGGGCACTGTGACATCTCTGAGGCTTCCGAGACA 479
QY 508 AGAGCTGTACCGGCTCAGACACACAGTGTGAGAGGGGCTGTGGGAGCTGGG 567
Db 480 AGAGCTGTACCGGCTCAGACACACAGTGTGAGAGGGGCTGTGGGAGCTGGG 539
QY 568 CGCTGCGGAGCGGGGAGCGCTCTCCGAGACCCCAAGCTCAGCTCAGCTGCTTCC 627
Db 540 CGCTGCGGAGCGGGGAGCGCTCTCCGAGACCCCAAGCTCAGCTCAGCTGCTTCC 599
QY 628 ACAGGAGATCTCGGGGCGAGAGGCTTCCAGAGCTGCAAGCTTCCGAGAGATGGCTGT 687
Db 600 ACAGGAGATCTCGGGGCGAGAGGCTTCCAGAGCTGCAAGCTTCCGAGAGATGGCTGT 659
QY 688 TCTGTGTGCGGAGTCTCGGCGGCGACCCCGGCGACTGCTGTGCTGTGCTTGGCC 747
Db 660 TCTGTGTGCGGAGTCTCGGCGGCGACCCCGGCGACTGCTGTGCTGTGCTTGGCC 719
QY 748 GCGAGCTGATCTCACTACCGGCTGTGCAAGCGGCGACCTTCAATGATGAGCGG 807
Db 720 GCGAGCTGATCTCACTACCGGCTGTGCAAGCGGCGACCTTCAATGATGAGCGG 779
QY 808 TGTCTCTCTGCAACCTATGGA -CATGTGAGAGCATTTACAGCA -GGACAAGGCGCTAT 865
Db 780 TGTCTCTCTGCAACCTATGGA -CATGTGAGAGCATTTACAGCA -GGACAAGGCGCTAT 839
QY 866 CTCGACCAAGCTGTGAGACCA 888
Db 840 CTCGACCAAGCTGTGAGACCA 862

RESULT 12
BI819535 912 bp mRNA linear EST 04-OCT-2001
LOCUS 60303674F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5177986.5,
DEFINITION mRNA sequence.
ACCESSION BI819535
VERSION BI819535.1 GI:15931085
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: sgabds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov

Plate: L14M1443 row: k column: 11
High quality sequence stop: 843.
Location/Qualifiers
1. 912
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5177986"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male brain, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."

BASE COUNT 153 a 310 c 299 g 150 t
ORIGIN

Query Match 38.1%; Score 762; DB 12; Length 912;
Best Local Similarity 96.1%; Pred. No. 1.6e-140;
Matches 867; Conservative 0; Mismatches 25; Indels 10; Gaps 8;

QY 1 CTCGCTCCAAAGTTGTGCAAGCCGGGAGCGCTCGGGGTGTGACGCGGCTCGGAGGCC 60
Db 15 CTCCTCTCAAGTTGTGCAAGCCGGGAGCGCTCGGGGTGTGACGCGGCTCGGAGGCC 74
QY 61 TCTGTGGGGGGGGGGGGGG--GGGGCTGCGGGGGGGGGGGGGGGGGGGGGGGGGGG 118
Db 75 TCTGTGG 134
QY 119 CAGGCTGCGGTGAGTGGGACCCAGCTCCTACCTCTGTGTCAGCGGCTGTGG 178
Db 135 CAGGCTGCGGTGAGTGGGACCCAGCTCCTACCTCTGTGTCAGCGGCTGTGG 194
QY 179 CAGGCTATTCG -AGCTGCCCGGAGTGTGACCACTGTGCTGCTGCTGCTGCTGCT 237
Db 195 CAGGCTATTCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 254
QY 238 CAGTTCCCTGCTGGGGGGGCGATGCGGGGGGAGGCTGTGCTGCTGCTGCTGCTGCT 297
Db 255 CAGTTCCCTGCT -GGGGGCGATGCGGGGGGAGGCTGTGCTGCTGCTGCTGCTGCT 313
QY 298 ACAGCTGTGATCTGTGAGGAATCTCCCGGTGAGCCCGGCTTCTCGAGCTGAGC 357
Db 314 ACAGCTGTGATCTGTGAGGAATCTCCCGGTGAGCCCGGCTTCTCGAGCTGAGC 373
QY 358 ACCCGCTCCGCTGTGACCGAGATGCCAAGAGGCGCTGGGCGCGGACCCAGTGT 417
Db 374 ACCCGCTCCGCTGTGACCGAGATGCCAAGAGGCGCTGGGCGCGGACCCAGTGT 433
QY 418 TCACCAATGTGAGACACACCGCGCCCAAGCGAGGGAGGCTGCTTCCGCAAGGGCGAG 477
Db 434 TCACCAATGTGAGACACACCGCGCCCAAGCGAGGGAGGCTGCTTCCGCAAGGGCGAG 492
QY 478 TGTGTACCATCTCTGAGGCGCTGCGAGAACAGAGCTGTACCGGTCAAGCACCACCA 537
Db 493 TGTGTACCATCTCTGAGGCGCTGCGAGAACAGAGCTGTACCGGTCAAGCACCACCA 552
QY 538 GTGACACAGAGGGGCTGTGCGAGCTGGGCGCTGCGGAGCGGGAGGCGCTTCCGAG 597
Db 553 GTGACACAGAGGGGCTGTGCGAGCTGGGCGCTGCGGAGCGGGAGGCGCTTCCGAG 612
QY 598 ACCCAAGCTCAGCTGTATGCGGTTCACAGGAGAGATCTCGGCGAGAGGCTGTCC 657
Db 613 ACCCAAGCTCAGCTGTATGCGGTTCACAGGAGAGATCTCGGCGAGAGGCTGTCC 672
QY 658 AGCAGCTGACGCTCCGAGAGATGGCTGTCTGCTGTGCGGAGATCCGCGGCCACCCCG 717

Db 673 AGCAGCTGCAGCTCCCGAGATGAGGCTGTTCTGTCGGGAGTCCG -GCGCCACCCCG 731
Qy 718 GCGACTACGCTCTGTCGCTGAGCTTTGGCGCGAGC -TCATCCACTACCGCGTGTGAC 776
Db 732 GCGACTACGCTCTGTCGCTGAGCTTTGGCGCGAGTTCATCCACTACCGCGTGTGAC 791
Qy 777 CCGGACGGCACCCTCACAATGATGAGCCGCTGTTCTTTCGCAACCTCATGACATGCTG 836
Db 792 -CGGACGGCACCCTCACAATGATGAGCCGCTGTTCTTTCGCAACCTCATGACATGCTG 849
Qy 837 GAGCATTTACA -GCAAGACAAAGGCGCTATCTGCACCAAGCTGTGAGACCAAGGCGGA 895
Db 850 GAGCATTTACAGCAGGAGGAGCGCGCATCTGCACCAAGCTGTGAGACCAAGGCGGA 909
Qy 896 AC 897
Db 910 AC 911
RESULT 13
BM925586 1014 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT 6710058 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5763792
DEFINITION 5' mRNA sequence.
ACCESSION BM925586
VERSION BM925586.1 GI:19375965
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL2816 row: d column: 01
High quality sequence start: 7
High quality sequence stop: 713.
Location/Qualifiers
1..1014
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5763792"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains; age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH-MGC library."

BASE COUNT 164 a 355 c 330 g 162 t 3 others
ORIGIN

Query Match 36.9%; Score 737.2; DB 12; Length 1014;
Best Local Similarity 96.0%; Pred. No. 1.4e-135;
Matches 842; Conservative 0; Mismatches 24; Indels 11; Gaps 8;
Qy 1 CCGGCTCCAAAGTTGTGACCGCGGACCGCTGCGGGGTGTGACCGCGGCTCGGAGAGCCC 60
|||||

Db 128 CTCGCTCCAAAGTTGTGACCGCGGACCGCTTCGGGGGTGTGACCGCGGCTCGGAGAGCCC 187
Qy 61 TCTCTGGGGGCGGGGCGCGG -GCGCTCTGGGGGCGCCCTTACGAGAAAGAGGAAGA 118
Db 188 TCTCTGGGGGCGGGGCGGGGCGGGCTCGGGGGCGCCCTTACGAGAAAGAGGAAGA 247
Qy 119 CCAGGCTGGTCCAGTGGCACCAGCCAGTCCCTACCTCTGTGTGCGAAGCGGCTGGCTGTG 178
Db 248 CCAGGCTGGTCCAGTGGCACCAGTCCCTACCTCTGTGTGCGAAGCGGCTGGCTGTG 307
Qy 179 CAGGCGATTCGACAGGCTCCGACGTGTGACCACTGTGAGTGTGCTTACCTGCTC 238
Db 308 CAGGCGATTCGACAGGCTCCGACGTGTGACCACTGTGAGTGTGCTTACCTGCTC 367
Qy 239 AGTTTCCTCTGTGGGGGCGATGGCGGGGCGAGGCTCTGTGTTCTCTGGCGGCAATTTC 298
Db 368 AGTTTCCTCTGT -GGGGGCGATGGCGGGGCGAGGCTCTGTGTTCTCTGGCGGCAATTTC 426
Qy 299 CGGCTGTGATTTCTGTGAGGAACCTTCCCGGGTGAAGCCCGCTTCTCCGAGCCTGGCA 358
Db 427 CGGCTGTGATTTCTGTGAGGAACCTTCCCGGGTGAAGCCCGCTTCTCCGAGCCTGGCA 486
Qy 359 CCGCCCTCCGCTCAGCCAGGATGCGAACGAGGCGCTGGGCGCGGAGCCAGCTGTAT 418
Db 487 CCGCCCTCCGCTCAGCCAGGATGCGAACGAGGCGCTGGGCGCGGAGCCAGCTGTAT 546
Qy 419 CACCAATGCGAGCACACCGCCCAAGCCAGGAGAGCTGCGCTTCCGCAAGGAGAGCT 478
Db 547 CACCAATGCGAGCACACCGCCCAAGCCAGGAGAGCTGCGCTTCCGCAAGGAGAGCT 606
Qy 479 GGTACACATCTGTGAGGCTGTGAGAGCAAGAGCTGTGACCGCTTCACACACACACAG 538
Db 607 GGTACACATCTGTGAGGCTGTGAGAGCAAGAGCTGTGACCGCTTCACACACACACAG 666
Qy 539 TGGACAGAGGAGGAGGCTGTGAGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 598
Db 667 TGGACAGAGGAGGAGGCTGTGAGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 726
Qy 599 CCGCAAGCTCAGCTCATGCTCCGTGTTCACAGGAGAGATCTCGGGGCAAGAGAGCTGTCA 658
Db 727 CCGCAAGCTCAGCTCATGCTCCGTGTTCACAGGAGAGATCTCGGGGCAAGAGAGCTGTCA 786
Qy 659 GCACTGCAAGCTCCCGAGAGATGGGCTGTCTGTGTGCGGAGTCCGGCGCACCCCGG 718
Db 787 GCACTGCAAGCTCCCGAGAGATGGGCTGTCTGTGTGCGGAGTCCGGCGCACCCCGG 846
Qy 719 GCACTGCAAGCTCCGTGTGAGGCTTT -GGCGGAGAGTCACTACACG -CGTGTGTCAC 776
Db 847 GCACTGCAAGCTCTGT 906
Qy 777 GCG -GACGCGCACCTCACAATGATG -AGCGCGTGTCTTCTGCAACCTCAT -GGACAT 832
Db 907 GCGGAGCGGCGCACCTCACAATGATGAGGCGGTGTCTTCTGCAAACTCATGTGGAGAT 966
Qy 833 GGTGGA -GCATTACAGCAAGGAGGAGGCGCTATCT 867
Db 967 GGTGAGACATTACCGCAAGGAGGAGGCGCTAT 1003
RESULT 14
BU617908 733 bp mRNA linear EST 23-SEP-2002
LOCUS BU617908/c
DEFINITION UI-H-DF0-bet-k-12-0-UI-s1 NCI-CGAP-DF0 Homo sapiens cDNA clone
UI-H-DF0-bet-k-12-0-UI 3', mRNA sequence.
ACCESSION BU617908
VERSION BU617908.1 GI:23284123
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 733)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcraps-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
PolyA-yes.

FEATURES
source Location/Qualifiers
1..733
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-bet-k-12-0-UI"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI-CGAP_DF0"
/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI-CGAP_DF0 is a cDNA library containing the following
tissue(s): Subchondral Bone. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GTTAGCGTC.
TAG_LIB=UI-H-DF0
TAG_TISSUE=Subchondral bone
TAG_SEQ=GTTAGCGTC"

BASE COUNT 100 a 248 c 226 g 158 t 1 others
ORIGIN

Query Match 35.4%; Score 707.6; DB 13; Length 733;
Best Local Similarity 99.2%; Pred. No. 8.8e-130;
Matches 721; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1274 CATGAGTACTCTGGAGAGCAAGAGCTTGTGACCGACCTGGCCCGCCGACATCTCT 1333
DB 733 CATGAGTACTCTGGAGAGCAAGAGCTTGTGACCGACCTGGCCCGCCGACATCTCT 675
QY 1334 GGTCTCAGAGAGCTGTGGCCAGAGTCAGGCACTTGTGCTGCAAGAGCCGAGCGAA 1393
DB 674 GGTCTCAGAGAGCTGTGGCCAGAGTCAGGCACTTGTGCTGCAAGAGCCGAGCGAA 615
QY 1394 GGGGCTAGACTCAAGCCGGCTGCCGTCAAGTGGAGCGGCCGAGGCTCTCAACAACGG 1453
DB 614 GGGGCTAGACTCAAGCCGGCTGCCGTCAAGTGGAGCGGCCGAGGCTCTCAACAACGG 555
QY 1454 GAAGTTACCAAGCAAGTCGATGTCGAGTTTGGGGTGTGCTGTGGAGGCTCTCTC 1513
DB 554 GAAGTTACCAAGCAAGTCGATGTCGAGTTTGGGGTGTGCTGTGGAGGCTCTCTC 495
QY 1514 ATATGACGAGGCTCTGACCTTAATAATGTCACTGAAGAGGTGTGAGAGCCGTGGAGAA 1573
DB 494 ATATGACGAGGCTCTGACCTTAATAATGTCACTGAAGAGGTGTGAGAGCCGTGGAGAA 435
QY 1574 GGGGTACCGCATGGAACCCCGGAGGCTGTCAAGGCCCGGTGACGCTCTCATGAGCAG 1633
DB 434 GGGGTACCGCATGGAACCCCGGAGGCTGTCAAGGCCCGGTGACGCTCTCATGAGCAG 375
QY 1634 CTGCTGGAGGCAAGCCCGCCGCGCCACCTTCGCGAAATGTCGCGAGAGCTGGC 1693
DB 374 CTGCTGGAGGCAAGCCCGCCGCGCCACCTTCGCGAAATGTCGCGAGAGCTGGC 315

QY 1694 CCGGAGTACGAGTGCAGAGTGCCTCCGCTCCAGAGGAGCCGAGCGGCTC 1753
DB 314 CCGGAGTACGAGTGCAGAGTGCCTCCGCTCCAGAGGAGCCGAGCGGCTC 255
QY 1754 CACCTGCCCCGAGAGCCAGAGAGCCCTGACCCACCCGCTGGGGCTTGGCCCCAGAGA 1813
DB 254 CACCTGCCCCGAGAGCCAGAGAGCCCTGACCCACCCGCTGGGGCTTGGCCCCAGAGA 195
QY 1814 CCGAGAGTGGAGAGTGGCGGCTGGGGGCACTGACAGGCCCCAAGAGAGGTCCAGCGG 1873
DB 194 CCGAGAGTGGAGAGTGGCGGCTGGGGGCACTGACAGGCCCCAAGAGAGGTCCAGCGG 135
QY 1874 GCAGTCATCTCTCTGTTGCCACAGCAGAGGGGCTGGCCAGTATGGGGGCTTGGGGCGC 1933
DB 134 GCAGTCATCTCTCTGTTGCCACAGCAGAGGGGCTGGCCAGTATGGGGGCTTGGGGCGC 75
QY 1934 CCGTGAACACCCGACCTGCGAAGATGATCGCCGATTAAGACGATTTAGAGACTC 1993
DB 74 CCGTGAACACCCGACCTGCGAAGATGATCGCCGATTAAGACGATTTAGAGAAAA 15
QY 1994 TAAAAA 2000
DB 14 AAAAAA 8

RESULT 15
BX369553
LOCUS
DEFINITION BX369553 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION BX369553
VERSION BX369553.1 GI:30457823
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 986)
AUTHORS L.H.W.B., Gruber,C., Jessee,J. and Polayres,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segre@genoscope.cns.fr; web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF0382F05.AF03591.2&cluster=9238.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue genoscope sequence ID : CS0BAF0382F05.AF03591_2.
Location/Qualifiers
1..986
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DU015Y001"
/cell_line="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 207 a 270 c 321 g 187 t 1 others
ORIGIN

Query Match 35.4%; Score 707; DB 13; Length 986;
Best Local Similarity 95.0%; Pred. No. 1.3e-129;

```
Matches 828; Conservative 0; Mismatches 31; Indels 13; Gaps 9;
QY 840 CATTACAGCAAGA-CAGGGGCGCTATCTGCACCAAGCTGTGTGACCAAAAGCGAACA 898
    |||
Db 3 CATTACAGCAAGAGCAAGGGCGCTATCTGCACCAAGCTGTGTGACCAAAAGCGAACA 62
QY 899 CGGGACCAAGTGGCCCGAGAGAGCTGGCCAGGGCGGCTGGTTACTGAACCTGCAGCA 958
    |||
Db 63 CGGGACCAAGTGGCCCGAGAGAGCTGGCCAGGGCGGCTGGTTACTGAACCTGCAGCA 122
QY 959 TTTGACATTGGGAGCAGATGGAGAGAGTTTGAGCTGCTGCAGGGTGAGTA 1018
    |||
Db 123 TTTGACATTGGGAGCAGATGGAGAGAGTTTGAGCTGCTGCAGGGTGAGTA 182
QY 1019 CCTGGGCAAAAGTGGCCGTGAAGATATCAAGTGTGATGTGACAGCCAGCCCTTCT 1078
    |||
Db 183 CCTGGGCAAAAGTGGCCGTGAAGATATCAAGTGTGATGTGACAGCCAGCCCTTCT 242
QY 1079 GGAACGACGCGCCGCTATGACGAGATGCAACAGAACTGTGTCTCTGGGCGT 1138
    |||
Db 243 GGAACGACGCGCCGCTATGACGAGATGCAACAGAACTGTGTCTCTGGGCGT 302
QY 1139 GATCCGACCAAGGGGCTGTACATGTCTGAGAGCAGTGAAGCAAGGCAACCTGTGAA 1198
    |||
Db 303 GATCCGACCAAGGGGCTGTACATGTCTGAGAGCAGTGAAGCAAGGCAACCTGTGAA 362
QY 1199 CTTTCTGCGGACCCGGGGTGAAGCCCTCGTGAACACCGCTCAGTCTCAGTTTCTCT 1258
    |||
Db 363 CTTTCTGCGGACCCGGGGTGAAGCCCTCGTGAACACCGCTCAGTCTCAGTTTCTCT 422
QY 1259 GCACGTGGCCGAGGGCATGAGTACCTGTGAGAGCAAGAACTTGTGCACCGGACCTGGC 1318
    |||
Db 423 GCACGTGGCCGAGGGCATGAGTACCTGTGAGAGCAAGAACTTGTGCACCGGACCTGGC 482
QY 1319 CGCCCGCAACATCTGTCTCAGAGAGACCTGGTGGCCAAAGTCAAGGCACTTGGGCTGGC 1378
    |||
Db 483 CGCCCGCAACATCTGTCTCAGAGAGACCTGGTGGCCAAAGTCAAGGCACTTGGGCTGGC 542
QY 1379 CAAAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGCTGCCGTCAAGTGGACGGCGCCGA 1438
    |||
Db 543 CAAAGCCGAGCGGAAGGGGCTAGACTCAAGCCGGCTGCCGTCAAGTGGACGGCGCCGA 602
QY 1439 GGCTCTCAACACAGGGAATTACCCAGCAAGTGGATGTCTGAG-TTTGGGGTGTCTGC 1497
    |||
Db 603 GGCTCTCAACACAGGGAATTACCCAGCAAGTGGATGTCTGAGTTTGGGGTGTCTGC 662
QY 1498 TCTGGAGGCTTCTCATATAT---GGAAGGGTCCGTAACCTTAAATGCTCAAGAAAGAG 1554
    |||
Db 663 TCTGGAGGAGTCTTCTCATATAGGAGCGGGCTCCGTAAATGTCACTGAAGAGAG 722
QY 1555 TGTGGAGGCG--GTGAGAGAGGGGTACCGCATGGAACCCCGAGGGC-TGTCCAGGCC 1611
    |||
Db 723 TGTGGAGGCGCGGTGAGATAGGGTATACCGCATGGAACCCCGAGGGCTTGTCCAGGCC 782
QY 1612 CCGTGCACGTCC-TGATGAGCAGC-TGCTGGGAGGCGAGA--GCCCGCCCGCGGCACCC 1667
    |||
Db 783 CCGGGACCTTCTTATGAGCAGCTTGTGGGAAGCAGAAAGCCCTGCCCGCGGCACCC 842
QY 1668 TTCCGCAACTGG-CCGAGAGACTGGCCCGGG 1698
    |||
Db 843 TTTCCGAAACTGTGTCGAGAAACTTGTCTCGG 874
```

Search completed: August 1, 2003, 19:41:12
Job time : 5081 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2003, 15:26:49 ; Search time 590 Seconds
(without alignments)
6993.223 Million cell updates/sec

Title: US-09-977-260-1
Perfect score: 2000
Sequence: 1 ctgcgtccaaagtgtgagc.....attctaagactcttaaaaa 2000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2679534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PC1US_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	2000	100.0	2000	9	US-09-977-269-1
2	2000	100.0	2000	10	US-09-977-260-1
3	2000	100.0	2000	11	US-09-977-261-1
4	1930	96.5	1989	14	US-10-084-817-341
5	547.6	27.4	2187	10	US-09-954-531-188
6	547.6	27.4	2187	15	US-10-298-377A-1
7	547.6	27.4	2420	14	US-10-177-293-87
8	192.6	9.6	2015	10	US-09-954-456-1983
9	192.6	9.6	2015	14	US-10-007-010-3
10	191.4	9.6	5527	10	US-09-880-107-3710
11	186.2	9.3	1911	10	US-09-917-800A-1611
12	185	9.2	1574	9	US-09-870-962-12
13	184.6	9.2	2354	10	US-09-967-768A-300
14	185.8	8.3	2451	10	US-09-771-161A-4
15	165.8	8.3	2674	13	US-10-003-295-1
16	162	8.1	2184	10	US-09-728-952-82

17	162	8.1	2520	11	US-09-953-047-3	Sequence 3, Appl1
18	160.4	8.0	2829	11	US-09-953-047-10	Sequence 10, Appl
19	155.6	7.8	2298	14	US-10-175-523-50	Sequence 50, Appl
20	148.8	7.4	2433	14	US-10-037-270-830	Sequence 830, App
21	148.8	7.4	2454	10	US-09-964-824A-120	Sequence 120, App
22	148.8	7.4	2598	14	US-10-072-036-110	Sequence 110, App
23	148.8	7.4	2616	14	US-10-072-036-108	Sequence 108, App
24	146.4	7.3	3454	13	US-10-044-090-48	Sequence 48, Appl
25	146.4	7.3	3726	9	US-09-925-702-271	Sequence 271, App
26	136.6	6.8	3945	9	US-09-921-771-41	Sequence 41, Appl
27	136.6	6.8	3969	10	US-09-982-610-23	Sequence 23, Appl
28	132.8	6.6	4698	10	US-09-895-652-1	Sequence 1, Appl1
29	131.6	6.6	3370	10	US-09-967-768A-144	Sequence 144, App
30	127.2	6.4	2710	14	US-10-084-817-161	Sequence 161, App
31	127.2	6.4	2820	10	US-09-924-859A-4	Sequence 4, Appl1
32	125.4	6.3	2301	14	US-10-242-943-3	Sequence 3, Appl1
33	124	6.2	4541	10	US-09-880-107-3785	Sequence 3785, App
34	124	6.2	4541	14	US-10-123-036-3	Sequence 3, Appl1
35	123.4	6.2	3914	13	US-10-044-090-148	Sequence 148, App
36	122	6.1	4221	14	US-10-220-801-8	Sequence 8, Appl1
37	122	6.1	6383	10	US-09-954-531-405	Sequence 405, App
38	121.2	6.1	2574	9	US-09-735-103-2	Sequence 2, Appl1
39	121.2	6.1	2574	13	US-10-045-428A-2	Sequence 2, Appl1
40	121.2	6.1	5264	10	US-09-920-300A-1731	Sequence 1731, App
41	121.2	6.1	5264	13	US-10-033-528-1731	Sequence 1731, App
42	120.8	6.0	3633	9	US-09-725-433-1	Sequence 1, Appl1
43	120.8	6.0	10058	9	US-09-974-298-98	Sequence 98, Appl
44	119.6	6.0	3416	8	US-08-987-689A-1	Sequence 1, Appl1
45	119.6	6.0	3416	14	US-10-292-524-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-977-269-1
Sequence 1, Application US/09977269
Patent No. US20020082037A1
GENERAL INFORMATION:
APPLICANT: ULBRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
APPLICANT: SURES, IRMINGARD
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1260
CURRENT APPLICATION NUMBER: US/09/977,269
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232,545
PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2000
TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
NAME/KEY: CDS
LOCATION: (258)..(1778)
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
OTHER INFORMATION: Kinase 1
US-09-977-269-1

Query Match 100.0%; Score 2000; DB 9; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTCGCTCCAAAGTTGGACGCGGCGCTCGGGGTGACAGCCGCTCGGAGGCC 60
DB 1 CTCGCTCCAAAGTTGGACGCGGCGCTCGGGGTGACAGCCGCTCGGAGGCC 60
OY 61 TCTGTGGGGGCGGCGGCGGCGCTCGGGGTGACAGCCGCTCGGAGGCC 120
DB 61 TCTGTGGGGGCGGCGGCGGCGCTCGGGGTGACAGCCGCTCGGAGGCC 120

QY 121 AGGCTGCTCAGTGGACACCAAGCTCCATCTCTGTGCGACCCGCTGGCTGTGGCA 180
 DB 121 AGGCTGCTCAGTGGACACCAAGCTCCATCTCTGTGCGACCCGCTGGCTGTGGCA 180
 QY 181 GGGCAATCCAGGTCGCCGACGCTGACCACTTGTCTCATGTGTCTCTACCTGCTCAG 240
 DB 181 GGGCAATCCAGGTCGCCGACGCTGACCACTTGTCTCATGTGTCTCTACCTGCTCAG 240
 QY 241 TTTCCCTCTGTGGGGGAGATGGGGGGGAGAGGCTCTGTGGTTCTGGGGGCAATTCAG 300
 DB 241 TTTCCCTCTGTGGGGGAGATGGGGGGGAGAGGCTCTGTGGTTCTGGGGGCAATTCAG 300
 QY 301 GCTGTATTTCTGTGAGAACTTCCCGGGGTGAGCCCCGCTTCTCCAGGCTGGACCC 360
 DB 301 GCTGTATTTCTGTGAGAACTTCCCGGGGTGAGCCCCGCTTCTCCAGGCTGGACCC 360
 QY 361 CCCCTCCGCTCAGGACGATGCCAAGAGGGGCTGGGGCCCCGGGGCACCAATGTATCA 420
 DB 361 CCCCTCCGCTCAGGACGATGCCAAGAGGGGCTGGGGCCCCGGGGCACCAATGTATCA 420
 QY 421 CCAAAATGCGAGACACCCGCCCAAGCAGGGAGCTGGCTTCCGCAAGGGGAGCTGG 480
 DB 421 CCAAAATGCGAGACACCCGCCCAAGCAGGGAGCTGGCTTCCGCAAGGGGAGCTGG 480
 QY 481 TCACCATCTGTGAGGCTTCCGAGAACAGAGCTGTACCGCGTCAAGCACACACAGTG 540
 DB 481 TCACCATCTGTGAGGCTTCCGAGAACAGAGCTGTACCGCGTCAAGCACACACAGTG 540
 QY 541 GACAGAGGGGGGCTGGGAGCTGGGGGCTGGGGAGGGAGGGAGGCTCTCCGAGACC 600
 DB 541 GACAGAGGGGGGCTGGGAGCTGGGGGCTGGGGAGGGAGGGAGGCTCTCCGAGACC 600
 QY 601 CCAAGCTCAGCTCATGCGGTGTTCCAGGGAGATCTCGGGCCAGAGGCTGTACAG 660
 DB 601 CCAAGCTCAGCTCATGCGGTGTTCCAGGGAGATCTCGGGCCAGAGGCTGTACAG 660
 QY 661 ACCTGACGCTTCCCGAGATGGGCTGTCTGTGCGGAGTCCGCGCCCAACCCGCGG 720
 DB 661 ACCTGACGCTTCCCGAGATGGGCTGTCTGTGCGGAGTCCGCGCCCAACCCGCGG 720
 QY 721 ACTAGTCTGTGCGGAGCTTGGCGCGGAGCTATCCACTACCGCGTGTCCACCGG 780
 DB 721 ACTAGTCTGTGCGGAGCTTGGCGCGGAGCTATCCACTACCGCGTGTCCACCGG 780
 QY 781 ACGGCCACCTCAATCGATGAGGCGGTCTTCTGTCAACCTCATGAGCATGTGAGAC 840
 DB 781 ACGGCCACCTCAATCGATGAGGCGGTCTTCTGTCAACCTCATGAGCATGTGAGAC 840
 QY 841 ATTACAGCAAGACAAAGGCGCTATCTGCAACCAAGCTGTGAGACCAAGGCAACAG 900
 DB 841 ATTACAGCAAGACAAAGGCGCTATCTGCAACCAAGCTGTGAGACCAAGGCAACAG 900
 QY 901 GGACCAAGTCCGCGGAGGAGAGCTGGGCCAGGGGCGGTGTTACTGAACTGTGCACATT 960
 DB 901 GGACCAAGTCCGCGGAGGAGAGCTGGGCCAGGGGCGGTGTTACTGAACTGTGCACATT 960
 QY 961 TGACATTGGGACACAGATCGAGAGGAGAGTTTGGAGCTCTCTGCAAGGCTGATACC 1020
 DB 961 TGACATTGGGACACAGATCGAGAGGAGAGTTTGGAGCTCTCTGCAAGGCTGATACC 1020
 QY 1021 TGGGGCAAAAGTGGCGGTGAAGAAATATCAATGTGTGATGACAGCCAGGCTTCTGG 1080
 DB 1021 TGGGGCAAAAGTGGCGGTGAAGAAATATCAATGTGTGATGACAGCCAGGCTTCTGG 1080
 QY 1081 ACGAGAGCGGCGTCAATGAGAGATGCAACAGAGAACTGTGCTCTCCGGGGGTGA 1140
 DB 1081 ACGAGAGCGGCGTCAATGAGAGATGCAACAGAGAACTGTGCTCTCCGGGGGTGA 1140
 QY 1141 TCCTGACACAGGGGCTGTATCTGTATGTGAGACAGTGAGCAAGGGCAACCTGTGACT 1200
 DB 1141 TCCTGACACAGGGGCTGTATCTGTATGTGAGACAGTGAGCAAGGGCAACCTGTGACT 1200

QY 1201 TTCTGCGGACCCGGGGTGAAGCCCTGTGTGAACACCGCTAGCTCTGCAATTTCTGTG 1260
 DB 1201 TTCTGCGGACCCGGGGTGAAGCCCTGTGTGAACACCGCTAGCTCTGCAATTTCTGTG 1260
 QY 1261 ACCTGGCCGAGGATGAGATTAAGTGAAGAGCAAGAGTGTGTGACACCGGACTGGCCG 1320
 DB 1261 ACCTGGCCGAGGATGAGATTAAGTGAAGAGCAAGAGTGTGTGACACCGGACTGGCCG 1320
 QY 1321 CCCGCAACATCTGTGCTCAGAGAGCTGGTGGCCCAAGGTCACAGCACTTGTGGCTGGCA 1380
 DB 1321 CCCGCAACATCTGTGCTCAGAGAGCTGGTGGCCCAAGGTCACAGCACTTGTGGCTGGCA 1380
 QY 1381 AACCCAGCGGAAAGGGCTAGACTCAAGCCGCTGCCCTCAAGTGAAGCGCCCGAGG 1440
 DB 1381 AACCCAGCGGAAAGGGCTAGACTCAAGCCGCTGCCCTCAAGTGAAGCGCCCGAGG 1440
 QY 1441 CTCTCAACACAGGGAAGTTACAGAGAGCGGATGTGTGAGATTTTGGGGTCTGCTG 1500
 DB 1441 CTCTCAACACAGGGAAGTTACAGAGAGCGGATGTGTGAGATTTTGGGGTCTGCTG 1500
 QY 1501 GGGAGGCTTCTCATATGACAGGGCTCCGTAACCTAAATGTCACTGAAGAGAGTGTGG 1560
 DB 1501 GGGAGGCTTCTCATATGACAGGGCTCCGTAACCTAAATGTCACTGAAGAGAGTGTGG 1560
 QY 1561 AGGCGCTGGAAGAGGGGTACCGCATGGAACCCCGAGAGGCTGTCCAGGCCCCGTACAG 1620
 DB 1561 AGGCGCTGGAAGAGGGGTACCGCATGGAACCCCGAGAGGCTGTCCAGGCCCCGTACAG 1620
 QY 1621 TCCTCATGAGAGCTGTGGGAGGAGAGCGGCGCGGCGGACCTTCCGAAACTGG 1680
 DB 1621 TCCTCATGAGAGCTGTGGGAGGAGAGCGGCGCGGCGGACCTTCCGAAACTGG 1680
 QY 1681 CCGAGAGAGCTGGCGCGGAGCTACGCAAGTGCACAGCTCCGTAACCTCTCAGGAGAG 1740
 DB 1681 CCGAGAGAGCTGGCGCGGAGCTACGCAAGTGCACAGCTCCGTAACCTCTCAGGAGAG 1740
 QY 1741 ACGCGAGGCTTCACCTGCGCCGGAAGCCAGAGGCTGTACCCCAACCGGTTGGGCTCT 1800
 DB 1741 ACGCGAGGCTTCACCTGCGCCGGAAGCCAGAGGCTGTACCCCAACCGGTTGGGCTCT 1800
 QY 1801 TGGCGCCAGAGGAGCGAGAGAGTGGAGGCTGGGGGACATGACAGGCCCAAG 1860
 DB 1801 TGGCGCCAGAGGAGCGAGAGAGTGGAGGCTGGGGGACATGACAGGCCCAAG 1860
 QY 1861 TGGCGCCAGAGGAGCGAGAGAGTGGAGGCTGGGGGACATGACAGGCCCAAG 1860
 DB 1861 TGGCGCCAGAGGAGCGAGAGAGTGGAGGCTGGGGGACATGACAGGCCCAAG 1860
 QY 1861 AGGCTCAGAGGAGGAGTATCTCTGCTGTGCGCCACAGAGGAGGCTGGCCACGTAGG 1920
 DB 1861 AGGCTCAGAGGAGGAGTATCTCTGCTGTGCGCCACAGAGGAGGCTGGCCACGTAGG 1920
 QY 1921 GGGCTGGGCGGCGCTGAGACACCCAGACCTGCGAAGGATGTCCCGCATTAAGACG 1980
 DB 1921 GGGCTGGGCGGCGCTGAGACACCCAGACCTGCGAAGGATGTCCCGCATTAAGACG 1980
 QY 1981 ATTCTAAGAGCTCTAATAAA 2000
 DB 1981 ATTCTAAGAGCTCTAATAAA 2000

RESULT 2
 US-09-977-260-1
 ; Sequence 1, Application US/09977260
 ; Publication No. US20020192790A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ULLRICH, AXEL
 ; APPLICANT: GISHIZKY, MIKHAIL
 ; APPLICANT: SURES, IRMINGARD
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
 ; FILE REFERENCE: 038602/1260
 ; CURRENT APPLICATION NUMBER: US/09/977, 260
 ; PRIOR APPLICATION NUMBER: 2001-10-16
 ; PRIOR FILING DATE: 2001-10-16
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1
LENGTH: 2000
TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
NAME/KEY: CDS
LOCATION: (258) ..(1778)
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
OTHER INFORMATION: kinase 1
US-09-977-260-1

Query Match 100.0%; Score 2000; DB 10; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGCTCCAAAGTTGTGCAACCGGACCGCTCGGGGTGTGCAAGCGGCTCGCGAAGCCC 60
DB 1 CTCGCTCCAAAGTTGTGCAACCGGACCGCTCGGGGTGTGCAAGCGGCTCGCGAAGCCC 60
QY 61 TCCTGGGGGCGGGGCGGGGCGGCTCGGGGCGGGCCCTGAGCAGAAAAGAGAAAC 120
DB 61 TCCTGGGGGCGGGGCGGGGCGGCTCGGGGCGGGCCCTGAGCAGAAAAGAGAAAC 120
QY 121 AGGCTCGGTGCAAGTGGACCCAGCTCCCTACCTCTGTGCAAGCGGCTGGCTGTGGCA 180
DB 121 AGGCTCGGTGCAAGTGGACCCAGCTCCCTACCTCTGTGCAAGCGGCTGGCTGTGGCA 180
QY 181 GGGCAATTCGCAAGCTGCGGACCTGTGACACTTGTCTCACTGTGCTTCACTGCTCAG 240
DB 181 GGGCAATTCGCAAGCTGCGGACCTGTGACACTTGTCTCACTGTGCTTCACTGCTCAG 240
QY 241 TTTCCTCTGGGGGCGATGGGGGCGAGGCTCTGTGTTCTGTTGCGGGGCAATTTCAG 300
DB 241 TTTCCTCTGGGGGCGATGGGGGCGAGGCTCTGTGTTCTGTTGCGGGGCAATTTCAG 300
QY 301 GCTGTGATTTCTGCTAGGAACCTTCCCGGGTGAGCCCGGCTTCTCCGAGCCTGACAC 360
DB 301 GCTGTGATTTCTGCTAGGAACCTTCCCGGGTGAGCCCGGCTTCTCCGAGCCTGACAC 360
QY 361 CCCCTCCGCTTCACACCCAGGATGCCACAGAGCGCTGGGCCCCGGGCAACCACTGTATCA 420
DB 361 CCCCTCCGCTTCACACCCAGGATGCCACAGAGCGCTGGGCCCCGGGCAACCACTGTATCA 420
QY 421 CCAATGCGAGCACACCGGCCCAAGCCAGGGAGGTGACCTTCCGCAAGGGCGAGCTG 480
DB 421 CCAATGCGAGCACACCGGCCCAAGCCAGGGAGGTGACCTTCCGCAAGGGCGAGCTG 480
QY 481 TCACCATCTGAGGCTGTGCGAACAAGAGCTGTACCGCTCAAGCAACCAACACAGTG 540
DB 481 TCACCATCTGAGGCTGTGCGAACAAGAGCTGTACCGCTCAAGCAACCAACACAGTG 540
QY 541 GACAGAGAGGGCTGTGAGAGCTGGGGGCGCTCGGGAGGGGAGGCCCTCTCTCCGAGACC 600
DB 541 GACAGAGAGGGCTGTGAGAGCTGGGGGCGCTCGGGAGGGGAGGCCCTCTCTCCGAGACC 600
QY 541 GACAGAGAGGGCTGTGAGAGCTGGGGGCGCTCGGGAGGGGAGGCCCTCTCTCCGAGACC 600
DB 541 GACAGAGAGGGCTGTGAGAGCTGGGGGCGCTCGGGAGGGGAGGCCCTCTCTCCGAGACC 600
QY 601 CCAAGCTAGCTCAATGCGGCTGTTCACAGGGAAGATCTCGGGCCAGAGAGGCTGTCCAC 660
DB 601 CCAAGCTAGCTCAATGCGGCTGTTCACAGGGAAGATCTCGGGCCAGAGAGGCTGTCCAC 660
QY 661 ACCTCAGGCTCCAGAGATGGGCTGTCTGTGGGGAGTCCGGGCGCCACCCCGGGG 720
DB 661 ACCTCAGGCTCCAGAGATGGGCTGTCTGTGGGGAGTCCGGGCGCCACCCCGGGG 720
QY 721 ACTAGTCTGTGCGTGAAGCTTTGGCCGGAGCTATCCACTACCGCGTGTGCAACCGG 780
DB 721 ACTAGTCTGTGCGTGAAGCTTTGGCCGGAGCTATCCACTACCGCGTGTGCAACCGG 780
QY 781 ACGGCACTCAATCAATGATGAGCGGCTGTCTTCACTCACTCAATGATGATGAGC 840
DB 781 ACGGCACTCAATCAATGATGAGCGGCTGTCTTCACTCACTCAATGATGATGAGC 840
QY 841 ATTACAGCAAGAGCAAGGGCGCTATCTGACCAAGCTGTGAGACCAGCAAGGCAACAG 900

DB 841 ATTACAGCAAGAGCAAGGGCGCTATCTGACCAAGCTGTGAGACCAGCAAGGCAACAGC 900
QY 901 GGACCAAGTCCGCCAGAGAGAGCTGCCAGGGCGGCTGTACTGAACCTGACACTT 960
DB 901 GGACCAAGTCCGCCAGAGAGAGCTGCCAGGGCGGCTGTACTGAACCTGACACTT 960
QY 961 TGACATTTGGAGCACAGATCGAGAGAGGAGATTGGAGCTGCTCCGCGAGGGTGAGTACC 1020
DB 961 TGACATTTGGAGCACAGATCGAGAGAGGAGATTGGAGCTGCTCCGCGAGGGTGAGTACC 1020
QY 1021 TGGGGCAAAAGGTGGCCGTGAAGATATCAAGTGTGATGACAGCCAGGCTTCTCTG 1080
DB 1021 TGGGGCAAAAGGTGGCCGTGAAGATATCAAGTGTGATGACAGCCAGGCTTCTCTG 1080
QY 1081 ACGAGACGGCCCTCATGACGAAGATGCAACACGAGAACCTGTGCTCTGCGGCTGA 1140
DB 1081 ACGAGACGGCCCTCATGACGAAGATGCAACACGAGAACCTGTGCTCTGCGGCTGA 1140
QY 1141 TCCTGACACAGGGGCTGTACATTGTATGAGACAGCTGAGCAAGGGCAACCTGTGAAC 1200
DB 1141 TCCTGACACAGGGGCTGTACATTGTATGAGACAGCTGAGCAAGGGCAACCTGTGAAC 1200
QY 1201 TTCTGCGGACCCGGGGTCAAGCCCTGTGTAACACCGCTACGCTCTGCAATTCTCTG 1260
DB 1201 TTCTGCGGACCCGGGGTCAAGCCCTGTGTAACACCGCTACGCTCTGCAATTCTCTG 1260
QY 1261 ACGTGGCGAGGGGCTGAGTACCTGTGAGAGCAAGAGCTTTGTGCAACCGGCTG 1320
DB 1261 ACGTGGCGAGGGGCTGAGTACCTGTGAGAGCAAGAGCTTTGTGCAACCGGCTG 1320
QY 1321 CCCGCAACATCTGCTGCTGAGAGGACCTGTGTCGCAAGGTCAGAGCACTTGGCTG 1380
DB 1321 CCCGCAACATCTGCTGCTGAGAGGACCTGTGTCGCAAGGTCAGAGCACTTGGCTG 1380
QY 1381 AAGCCGAGGGAAGGGGCTAGACTCAAGCCGGCTGCCCTCAAGTGTGAGCGGCCGAG 1440
DB 1381 AAGCCGAGGGAAGGGGCTAGACTCAAGCCGGCTGCCCTCAAGTGTGAGCGGCCGAG 1440
QY 1441 CTCTCAACACGGGAAGTTCAACAGCAAGTCCGAGTGTGTGAGTTTGGGGTGTCT 1500
DB 1441 CTCTCAACACGGGAAGTTCAACAGCAAGTCCGAGTGTGTGAGTTTGGGGTGTCT 1500
QY 1501 GGGAGGCTTCTCATATGAGAGGGGCTCCCTACCTCAAAATGTACATGAAGAGTGTG 1560
DB 1501 GGGAGGCTTCTCATATGAGAGGGGCTCCCTACCTCAAAATGTACATGAAGAGTGTG 1560
QY 1561 AGGCGGTGAGAGGGGTAACCGATGAACCCCGAGGGCTGTCCAGGCCCTGTGACG 1620
DB 1561 AGGCGGTGAGAGGGGTAACCGATGAACCCCGAGGGCTGTCCAGGGCTGTGACG 1620
QY 1621 TCTCATGAGCAAGCTGTGAGAGGAGAGAGCCCGGCGGCGGCTTCCGCAACTG 1680
DB 1621 TCTCATGAGCAAGCTGTGAGAGGAGAGAGCCCGGCGGCGGCTTCCGCAACTG 1680
QY 1681 CCGAGAAGCTGAGCCGGAGCTACCGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1740
DB 1681 CCGAGAAGCTGAGCCGGAGCTACCGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1740
QY 1741 ACGCGGAGGCTCAACCTGCGCCGAAGCGAGAGGCTGACCCCAACCGGTTGGGGCT 1800
DB 1741 ACGCGGAGGCTCAACCTGCGCCGAAGCGAGAGGCTGACCCCAACCGGTTGGGGCT 1800
QY 1801 TGGGCCGAGAGAGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1860
DB 1801 TGGGCCGAGAGAGCGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1860
QY 1861 AAGGTCAGAGGCGGGAAGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
DB 1861 AAGGTCAGAGGCGGGAAGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
QY 1921 GGGTGTGGGGCGGCGGTGAGCAACCCAGACCTGCGAAGGATGCTGCCGATTAAGACG 1980

Db 1921 GGCCTGGGGCGCCCTGGACACCCAGACCTGCGAAGATGATCGCCGATAAAGACGG 1980
QY 1981 ATTCTAAGACTCTAAAAA 2000
Db 1981 ATTCTAAGACTCTAAAAA 2000

RESULT 3

US-09-977-261-1
; Sequence 1, Application US/09977261
; Publication No. US20030054527A1
; GENERAL INFORMATION:
; APPLICANT: ULIRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAEL
; APPLICANT: SURES, IMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1259
; CURRENT APPLICATION NUMBER: US/09/977, 261
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (258)..(1778)
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 1
US-09-977-261-1

Query Match: 100.0%; Score 2000; DB 11; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCAAGATTGTGTCAGCGCGGACCGCTCGGGGTGTGTCAGCGCGCTCGCGAGAGCC 60
Db 1 CTGCTCAAGATTGTGTCAGCGCGGACCGCTCGGGGTGTGTCAGCGCGCTCGCGAGAGCC 60
QY 61 TCTTGAGGGGCGGCGGCGGCGGCTGCGGGGCGCCCTGTGAGCAAGAAAGAGAAC 120
Db 61 TCTTGAGGGGCGGCGGCGGCGGCTGCGGGGCGCCCTGTGAGCAAGAAAGAGAAC 120
QY 121 AGGCTGGTCCATGAGCAACCCAGCTCCCTACCTGCTGCTGTCAGCGCGCTGCTGCGCA 180
Db 121 AGGCTGGTCCATGAGCAACCCAGCTCCCTACCTGCTGCTGTCAGCGCGCTGCTGCGCA 180
QY 181 GGCATTTCCAGCGTCCCGACTGTGACACTTGTCTAGTGTCTCTCACTGCTCAG 240
Db 181 GGCATTTCCAGCGTCCCGACTGTGACACTTGTCTAGTGTCTCTCACTGCTCAG 240
QY 241 TTTCTCTGCGGGGCGATGCGGCGGCGAGGCTCTCTGTTTCTGCGGGGCAATTTACG 300
Db 241 TTTCTCTGCGGGGCGATGCGGCGGCGAGGCTCTCTGTTTCTGCGGGGCAATTTACG 300
QY 301 GCGTATTTCTGCTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCAGCTTGGCACC 360
Db 301 GCGTATTTCTGCTGAGAACTTCCCGGGGTGAGCCCGCTTCTCCAGCTTGGCACC 360
QY 361 CCCCTCCGCTCAGCAGAGATGCAAGAGAGCGCTGGGGCCCGGCGACCACTGATCA 420
Db 361 CCCCTCCGCTCAGCAGAGATGCAAGAGAGCGCTGGGGCCCGGCGACCACTGATCA 420
QY 421 CCAATGCGAGCAGACCCGCCCAAGCAGGGGAGGTGGCTTCCCAAGGGGAGAGTGG 480
Db 421 CCAATGCGAGCAGACCCGCCCAAGCAGGGGAGGTGGCTTCCCAAGGGGAGAGTGG 480
QY 481 TCACCATCTGAGGCGCTGCGAGAAACAAGCTGTGTAACCGCTCAAGCAGACCACTG 540
Db 481 TCACCATCTGAGGCGCTGCGAGAAACAAGCTGTGTAACCGCTCAAGCAGACCACTG 540

Db 481 TCACCATCTGAGGCGCTGCGAGAAACAAGCTGTGTAACCGCTCAAGCAGACCACTG 540
QY 541 GACAGAGGGGCTGTGTCAGCTGGGGCGTGGCGGAGCGGAGAGCCCTCTCCGACAGC 600
Db 541 GACAGAGGGGCTGTGTCAGCTGGGGCGTGGCGGAGCGGAGAGCCCTCTCCGACAGC 600
QY 601 CCAAGCTCAGCTCATGCGCTGTGTTCCACGGGAAGATCTGGGGCCAGAGAGCTGTCCAGC 660
Db 601 CCAAGCTCAGCTCATGCGCTGTGTTCCACGGGAAGATCTGGGGCCAGAGAGCTGTCCAGC 660
QY 661 AGCTGCAAGCTCCCGAGAGATGGGCTGTCTGTTGCGGGAGTCCCGCGCCACCCCGCG 720
Db 661 AGCTGCAAGCTCCCGAGAGATGGGCTGTCTGTTGCGGGAGTCCCGCGCCACCCCGCG 720
QY 721 ACTACGCTGTGCTGAGCTTTGGCGGCGACGTCACCTACCGGCTGTGACCGCG 780
Db 721 ACTACGCTGTGCTGAGCTTTGGCGGCGACGTCACCTACCGGCTGTGACCGCG 780
QY 781 ACGGCCACCTCAGCAATGATGAGGCGCTGTTCTTCTGCAACCTCATGAGATGGTGGAGC 840
Db 781 ACGGCCACCTCAGCAATGATGAGGCGCTGTTCTTCTGCAACCTCATGAGATGGTGGAGC 840
QY 841 ATTACAGCAAGGACAGAGGCGCTATCTGCAACAGCTGTGTGAGACCAACCGAAGACAG 900
Db 841 ATTACAGCAAGGACAGAGGCGCTATCTGCAACAGCTGTGTGAGACCAACCGAAGACAG 900
QY 901 GGAACCAAGTGGGCGGAGAGAGCTGGCGAGGGCGGCTGTGTTACTGTAACCTGACAGATT 960
Db 901 GGAACCAAGTGGGCGGAGAGAGCTGGCGAGGGCGGCTGTGTTACTGTAACCTGACAGATT 960
QY 961 TGACATTGGGAGCAGAGATGCGAGAGGAGAGTGGAGCTGTCCTGACAGGTTGATACC 1020
Db 961 TGACATTGGGAGCAGAGATGCGAGAGGAGAGTGGAGCTGTCCTGACAGGTTGATACC 1020
QY 1021 TGGGGCAAAAGTGGCCGCTGAAGAATATCAAGTGTGATGTGAGACGCCAGCTTCTG 1080
Db 1021 TGGGGCAAAAGTGGCCGCTGAAGAATATCAAGTGTGATGTGAGACGCCAGCTTCTG 1080
QY 1081 ACAGAGAGCGGCTCATGACGAAGATGCAACACAGAACTGCTGCTGCTGGCGTGA 1140
Db 1081 ACAGAGAGCGGCTCATGACGAAGATGCAACACAGAACTGCTGCTGCTGGCGTGA 1140
QY 1141 TCTGCAACAGGGGCTGTATCTATGTGAGAGCAGTGTGAGCAAGGGCACTGGTGA 1200
Db 1141 TCTGCAACAGGGGCTGTATCTATGTGAGAGCAGTGTGAGCAAGGGCACTGGTGA 1200
QY 1201 TTTCTGGGAGCCGCGGGGTGCGGCCCTGTGTAACCCGCTGACGCTCTGCAATTTCTG 1260
Db 1201 TTTCTGGGAGCCGCGGGGTGCGGCCCTGTGTAACCCGCTGACGCTCTGCAATTTCTG 1260
QY 1261 ACCTGGCCGAGGAGATGAGTACCTGTGAGAGCAAGAACTTGTGCAACCGGACCTGGCG 1320
Db 1261 ACCTGGCCGAGGAGATGAGTACCTGTGAGAGCAAGAACTTGTGCAACCGGACCTGGCG 1320
QY 1321 CCCGCAACATCTGTGCTCAGAGAGCTGTGTCGCAAGGTGACGACTTTGGCTGGCCA 1380
Db 1321 CCCGCAACATCTGTGCTCAGAGAGCTGTGTCGCAAGGTGACGACTTTGGCTGGCCA 1380
QY 1381 AAGCCGAGGGAAGGGGCTGAGTCAAGCCGCGTGCCTGCTGAGTGGAGCGGCGCGAGG 1440
Db 1381 AAGCCGAGGGAAGGGGCTGAGTCAAGCCGCGTGCCTGCTGAGTGGAGCGGCGCGAGG 1440
QY 1441 CTCTCAACACAGGGAAGTTCACAGAGAAGTGTGATGTTTGGGGTCTGCTCT 1500
Db 1441 CTCTCAACACAGGGAAGTTCACAGAGAAGTGTGATGTTTGGGGTCTGCTCT 1500
QY 1501 GGGAGGCTTCTCATATGAGAGGCTCCGTAACCTTAATGTCACTGAAAGAGTGTGG 1560
Db 1501 GGGAGGCTTCTCATATGAGAGGCTCCGTAACCTTAATGTCACTGAAAGAGTGTGG 1560
QY 1561 AGGCGGTGGAAGAGGGGTACCGCATGGAACCCCGGAGGGCTTCCAGGCGCGCTGACAG 1620
Db 1561 AGGCGGTGGAAGAGGGGTACCGCATGGAACCCCGGAGGGCTTCCAGGCGCGCTGACAG 1620

Db 1260 GCACTGGCCGAGGCGATGAGTACCTGTGAGAGCAAGAACTTGTGACACCGGACCTGGC 1319
QY 1319 CGCCGCAACATCTGTGTCTCAGAGAGACCTGTGTGCGCAAGTACGACCTTTGGCTGGC 1378
Db 1320 CGCCGCAACATCTGTGTCTCAGAGAGACCTGTGTGCGCAAGTACGACCTTTGGCTGGC 1379
QY 1379 CAAAGCGAGCGAAGGAGGCTAGACTCAAGCCGGCTGCCGCTCAAGTACGAGCGCGCCGA 1438
Db 1380 CAAAGCGAGCGAAGGAGGCTAGACTCAAGCCGGCTGCCGCTCAAGTACGAGCGCGCCGA 1439
QY 1439 GGCCTCAACAACGAGAACTTACCAAGCAAGTGTCTGTGAGAGTTTGGGTGTGCT 1498
Db 1440 GGCCTCAACAACGAGAACTTACCAAGCAAGTGTCTGTGAGAGTTTGGGTGTGCT 1499
QY 1499 CTGGGAGTCTTCTCATATGAGAGGCGCTCGTACCTTAATAATGTCTACTGAAGAGGTGC 1558
Db 1500 CTGGGAGTCTTCTCATATGAGAGGCGCTCGTACCTTAATAATGTCTACTGAAGAGGTGC 1559
QY 1559 GGAAGCGGTGGAAGAGGCTACCGCATGGAACCCCGGAGGCGTGTCCAGGCCCGCTGCA 1618
Db 1560 GGAAGCGGTGGAAGAGGCTACCGCATGGAACCCCGGAGGCGTGTCCAGGCCCGCTGCA 1619
QY 1619 GTCCTCATAGAGAGTGTGTGGAGGAGAGCGCCCGCGCCGCAACCTTCCGCAACT 1678
Db 1620 GTCCTCATAGAGAGTGTGTGGAGGAGAGCGCCCGCGCGCAACCTTCCGCAACT 1678
QY 1679 GGCAGGAAGTGTGGCCGCGGAGCTACGAGTGCAGTGTGCCAGCTTCCGTCAAGGCA 1738
Db 1679 GGCAGGAAGTGTGGCCGCGGAGCTACGAGTGCAGTGTGCCAGCTTCCGTCAAGGCA 1738
QY 1739 GGAAGCGGAGT 1797
Db 1739 GGAAGCGGAGT 1797
QY 1798 CTTGGCCCGCAGAGAGCGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1857
Db 1798 CTTGGCCCGCAGAGAGCGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1857
QY 1858 AGGAGGCTCAGAGCGGAGGAGTATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1917
Db 1858 AGGAGGCTCAGAGCGGAGGAGTATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1917
QY 1918 GGGGGCTCTGGGCGCGCTGTGACACCCGAGCTGTGGAAGATGTATCCCGCATTAAGA 1977
Db 1918 GGGGGCTCTGGGCGCGCTGTGACACCCGAGCTGTGGAAGATGTATCCCGCATTAAGA 1977
QY 1978 CGGATTCTAAG 1989
Db 1978 CGGATTCTAAG 1989

RESULT 5
US-09-954-531-188
; Sequence 188, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392

SOFTWARE: PatentIn version 3.0
; SEQ ID NO 188
; LENGTH: 2187
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-188

Query Match
Best Local Similarity 64.9%; Pred. No. 3.3e-121;
Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

27.4%; Score 547.6; DB 10; Length 2187;
Query Match
Best Local Similarity 64.9%; Pred. No. 3.3e-121;
Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

QY 395 CTGGGCGCGGCGGCGGAGTGTATCAACCAATGCGAGCACACCGCCCAAGCGAGGGA 454
Db 154 CTGGGCGCGGCGGCGGAGTGTATTCAGCAACTTCCAGGCGAGTCCGCGAGGCA 213
QY 455 GGTGGCTTCCGCAAGGAGGAGTGTGTACATCTCCGAGGCGCTGGCGCAAGAGCTG 514
Db 214 CTTGCCCTTGTCAAGAGGAGAGCTGTACCATTTGTGGCGCTGACCAAGAGCCCACTG 273
QY 515 GTACCGCTCAAGACACACACAGTGTGAGAGAGGAGGCTGTGTGCACTGTGGGCGCTGC 574
Db 274 GTACCAAGGCA--AAACAGGTGGCGGTGAGGCGATCATGCCAAGTACTGCA 330
QY 575 GGAAGCGGAGGCGGCTCTCGGAGACCCCAAGCTCAAGCTCATGTGCTGTCCAGCGGA 634
Db 331 GAAAGCGGAGGCGGTGTGAAAGGCGGTACCAAACTCACTCATATGCTTGTCCAGCGCA 390
QY 635 GATCTGCGGCGAGGAGTGTGTGAGAGCTGTGAGAGCTGTGTGAGTGTGTGTGTGTGTGT 694
Db 391 GATCTGCGGAGGAGTGTGTGAGAGCTGTGTGAGAGCTGTGTGAGTGTGTGTGTGTGTGT 450
QY 695 GCGGAGTGTGCGGCGGCGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 754
Db 451 GCGGAGTGTGCGGCGGCGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 510
QY 755 CATCCATCTCCGCTGT 814
Db 511 GGAAGCTACCTGATATGTATGTATGTATGTATGTATGTATGTATGTATGTATGTATGT 570
QY 815 CTGCACTCATGAGATGT 874
Db 571 TGAGAACCTCATGAGT 630
QY 875 GCTGTGTGAGACCAAGAGCGGAAACGCGGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGT 934
Db 631 CCTCATTAACCAAGAGT 690
QY 935 GGGCTGT 994
Db 691 CGGCTGT 750
QY 995 TGGAGCTGTCTGAGAGGAGT 1054
Db 751 CGGAGAGT 810
QY 1055 TGTGTGTGACAGCCCGAGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1114
Db 811 CGAGCGCATGTGCGGAGGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 870
QY 1115 GAACGT 1168
Db 871 CAACCTGT 930
QY 1169 GGAAGCTGTGAGAGGAGGAGT 1228
Db 931 TGAGTACATGAGGAGGAGGAGT 990
QY 1229 GAAACAGCTGT 1288
Db 991 GGGCGGAGAGT 1050
QY 1289 GAGCAAGAGCTGT 1348

Db 1051 GGGCAACAATTTGTCGATCGACAGACCTGGCTGCCCGCAATGTGCTGTGAGACAA 1110
Qy 1349 GGTGGCCAAAGTTCAGACGACTTTGGCTGGCCAAAGCCGAGGGAAGGGGCTAGACTCAAG 1408
Db 1111 CGTGGCCAAAGTTCAGACGACTTTGGCTGGCCAAAGCCGAGGGAAGGGGCTAGACTCAAG 1170
Qy 1409 CGGCTGGCCCTCAAGTGGACGGCGCCCGGAGGCTTCAACACGCGGAATTCACACAGCA 1468
Db 1171 CAAGTGGCCCTCAAGTGGACGGCGCCCGGAGGCTTCAACACGCGGAATTCACACAGCA 1230
Qy 1469 GTCGATGCTGAGATTTTGGGGTGTCTCTGGAGAGTCTTCTATATGACAGGGCTCC 1528
Db 1231 GTCGATGCTGAGATTTTGGGGTGTCTCTGGAGAGTCTTCTATATGACAGGGCTCC 1290
Qy 1529 GTACCTTAAATGTCTACTGAAAGAGTGTGAGAGCCGTGAGAGGGGTACCCGATGGA 1588
Db 1291 TTATCCAAAGATTTCCCTCGAAGAGCTGTCCCTCGGTGGAGAAAGGCTTACAGATGGA 1350
Qy 1589 ACCCCCGAGGGCTGTCCAGGCCCCGTGACGCTCTCATATGAGCAGCTGTGGAGGAGCA 1648
Db 1351 TGCCCCGAGGGCTGTCCAGGCCCCGTGACGCTCTCATATGAGCAGCTGTGGAGGAGCA 1410
Qy 1649 GCCCGCCCGGCGCCACCTTCCGCAAACTGGCCGAGAGCTGG 1692
Db 1411 GCGCGCCATGCGGGCCCTCTCTCTACAGCTCCGAGAGCAGCTTG 1454

RESULT 6

US-10-298-377A-1
; Sequence 1, Application US/10298377A
; Publication No. US20030130209A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Cheresif, David A.
; APPLICANT: Paul, Robert
; APPLICANT: Elicevli, Brian
; TITLE OF INVENTION: Method of Treatment of Myocardial
; FILE REFERENCE: TSRI-651.5
; CURRENT APPLICATION NUMBER: US/10/298,377A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 10/298,377
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/470,881
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 09/538,248
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: PCT/US99/11780
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/087,220
; PRIOR FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2187
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)...(1486)
US-10-298-377A-1

Query Match 27.4%; Score 547.6; DB 15; Length 2187;
Best Local Similarity 64.9%; Pred No. 3.3e-121;
Matches 846; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

Qy 395 CTGGGCCCCGGGACCCAGTATACCAAAATGCGAGACACCCGCCCAAGCCAGGGGA 454
Db 154 CTGGCATTCGGTACAGATGATGCAAGTAACTTCCACGCGCACTGCCGACAGGA 213
Qy 455 GCTGGCCTTCGCAAGGCGAGCTGTGACCATCTCTGAGAGCCTTCGGAAGACAGAGCTG 514
Db 214 CTTGCCCTTTCGCAAGGAGAGAGCTGTGACCATCTGTGCGCCGACCAAGAGCCCACTG 273

Qy 515 GTACCGCTCAAGCACACACAGTGCACAGAGGGGCTGTGCGAGCTGGGGCGTCG 574
Db 274 GTACCAAGCCA---AAACAAAGTGGCGCGTGAAGGACTCATCCAGCCAACTAGTCCA 330
Qy 575 GGAGGGGAGGCGCTCTCCGAGACCCCAAGCTCAGCTTATGCCGTGGTTCCAGGGGA 634
Db 331 GAAGGGGAGGCGGTGAAGGGGAGTCAACAACTAGCTTATGCCCTTGGTTCCAGGGGA 390
Qy 635 GATCTGGGCGCAGAGGCTGTCCAGCAGCTCAGCTCCAGAGATGGGCTGTCTCTG 694
Db 391 GATCAGAGGAGAGCGGTGAGCGCTTGTATCCCGGAGAGACAGGCGCTTCTCTG 450
Qy 695 GCGGAGTCCCGCGGCCACCCCGGAGCTAGCTGCTGCGCTGAGACTTTGGCCGAGAGT 754
Db 451 GCGGAGAGACCAACTACCCCGAGACTTACAGCTGTGCTGAGCTCCGAGCGAGAGT 510
Qy 755 CATCCATACCGGCTGTGACACCGGAGCGCCACTCAATTCGATGAGGCCGTGTTCT 814
Db 511 GGAGCATACCGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 570
Qy 815 CTGCACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 874
Db 571 TGAGAACCTCATGAGCTGTGAGACACTACACTACAGCAGATGAGACTGTATACCG 630
Qy 875 GCTGTGAGACCAAGCGGAAACAGGGAGCCAGCGAGGAGGAGGAGGAGGAGGAGG 934
Db 631 CCTCATTAACCAAGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 690
Qy 935 GGGCTGTACTGAACTGACAGATTTGACATTTGGAGCAGACAGATCGAGAGGAGAGT 994
Db 691 CGGCTGGCCCTGAACATGAAAGAGAGCTGAGCTGTGAGACATCGGAGGAGGAGGAGT 750
Qy 995 TGGAGCTTCTGCAAGGCTGAGTACTTGGGCAAAAGTGGCCGTGAAGATATCAAGT 1054
Db 751 CGGAGACTGATGCTGGGCGATTAACGAGGAGCAAACTCCCGCAAGTCAATTAAGAA 810
Qy 1055 TGATGTGACAGCCCGGCTTCTGAGAGAGAGGCGCCCTGATGAGAGATGCAACAGA 1114
Db 811 CGAGCCCACTGCGCCAGGCTTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 870
Qy 1115 GAACCTGTGCTCTCTGAGGCGTGTATCTGACAG-----GGCTGTACATTTGAT 1168
Db 871 CAACCTGTGAGAGCTCTGGCGGTGTGTGAGAGAGAGGAGGAGGAGGAGGAGGAGG 930
Qy 1169 GGAGCAGCTGAGCAAGGCAACCTGTGATCTTCTGAGAGCCGGGGTCCAGCCCTCTG 1228
Db 931 TGAGTACATGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 990
Qy 1229 GAACACCGCTCAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1288
Db 991 GGGCGGAGAGCTGTCTCTCAAGTTCTCTGATGATGATGATGATGATGATGATGATG 1050
Qy 1289 GAGCAAGAAGCTGTGACACCGGAGCTGTGCGCCGCAACATCTGTCTGAGAGAGCT 1348
Db 1051 GGGCAACAATTTGCGATCGACAGCTGTGCGCCGCAAAATGCTGTGCTGAGAGCAA 1110
Qy 1349 GGTGGCCAAAGTTCAGACGACTTTGGCTGGCCAAAGCCGAGGGAAGGGGCTAGACTCAAG 1408
Db 1111 CGTGGCCAAAGTTCAGACGACTTTGGCTGGCCAAAGCCGAGGAGGAGGAGGAGGAG 1170
Qy 1409 CGGCTGGCCCTCAAGTGGACGGCGCCCGGAGGCTTCAACACGCGGAATTCACACAGCA 1468
Db 1171 CAAGTGGCCCTCAAGTGGACGGCGCCCGGAGGCTTCAACACGCGGAATTCACACAGCA 1230
Qy 1469 GTCGATGCTGAGATTTTGGGGTGTCTCTGGAGAGTCTTCTATATGACAGGGCTCC 1528
Db 1231 GTCGATGCTGAGATTTTGGGGTGTCTCTGGAGAGTCTTCTATATGACAGGGCTCC 1290
Qy 1529 GTACCTTAAATGTCTACTGAAAGAGTGTGAGAGCCGTGAGAGAGGGGTACAGATGGA 1588
Db 1291 TTATCCAAAGATTTCCCTCGAAGAGGAGTGTGCTCTGCGGTGAGAGAGGCGCTACAGATGGA 1350

RESULT 14
US-09-771-161A-4

; Sequence 4, Application US/09771161A
; Patent No. US2002010811A1

; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A

; PRIORITY FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: -
; LOCATION: (1)..(2451)
; OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'
US-09-771-161A-4

Query Match 8.3%; Score 165.8; DB 10; Length 2451;
Best Local Similarity 56.0%; Pred. No. 3.3e-30;
Matches 361; Conservative 0; Mismatches 272; Indels 12; Gaps 2;

QY 1060 TGACGCCAGCCGCTTCCTGCGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1119
DB 1262 TGTCTCCAGAACCCCTTCCTGCGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1321
QY 1120 TGTGCTGCTCTGCGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1179
DB 1322 TGTGCTGCTCTGCGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1381
QY 1180 GCAAGGCAACCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1239
DB 1382 CCAGAGGATGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1441
QY 1240 AGCTCTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1299
DB 1442 GGTCTATGATGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1501
QY 1300 TTGTCACCCGCGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1359
DB 1502 CCATCCACCGCGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1561
QY 1360 TCAGGACTTGGCTGAGC-----CAAGCCGAGAGAGAGAGCGCCGCTCATAGCAAGC 1410
DB 1562 TTGCTGATTTGGCTGAGC-----CAAGCCGAGAGAGAGAGCGCCGCTCATAGCAAGC 1621
QY 1411 GGTCTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1470
DB 1622 AGTCCCATCAAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1681
QY 1471 GGTCTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1530
DB 1682 CAGAGAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1741
QY 1531 ACCCTAAATGCTCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1590
DB 1742 ACCCTAAATGCTCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1801
QY 1591 GGTCTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1647
DB 1802 GGTCTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1861

QY 1648 AGCCCGCCCGCCGCGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1692
DB 1862 GGTCTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1906

RESULT 15
US-10-003-295-1

; Sequence 1, Application US/10003295
; Publication No. US20020168741A1

; GENERAL INFORMATION:
; APPLICANT: GAN, Weimin et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; PRIORITY FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows version 4.0

; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-1

Query Match 8.3%; Score 165.8; DB 13; Length 2674;
Best Local Similarity 55.6%; Pred. No. 3.4e-30;
Matches 370; Conservative 0; Mismatches 277; Indels 18; Gaps 2;

QY 1042 AGAATATCAAGTGTGATGATGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1101
DB 1642 AGAGCTCCAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1701
QY 1102 AGATCAACAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1161
DB 1702 AGAGCTCCAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1761
QY 1162 TGTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1221
DB 1762 TGTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1821
QY 1222 CCCTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1281
DB 1822 CCCTGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1878
QY 1282 ACCTGAGAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1341
DB 1878 ACCTGAGAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1938
QY 1342 AGAGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1401
DB 1938 AGAGCTGCTGCTGCAAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1998
QY 1402 ACTCA-----AGCCGCGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1446
DB 1998 ACTCA-----AGCCGCGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 2058
QY 1447 AACAGAGAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1506
DB 2058 AACAGAGAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 2118
QY 1507 TCTTCTCATAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1566
DB 2118 TCTTCTCATAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 2178
QY 1567 TGGAGAGAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1626
DB 2178 TGGAGAGAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 2238
QY 1627 TGGAGAGAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 1686
DB 2238 TGGAGAGAGAGAGAGAGAGAGAGCGCCGCTCATAGCAAGATGCAACGAGAAC 2298

Qy	1687	AGCTG	1691
Db	2299	AGCTG	2303

Search completed: August 1, 2003, 19:54:25
Job time : 597 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2003, 19:41:19 ; Search time 3240 Seconds
(without alignments)
3803.200 Million cell updates/sec

Title: US-09-977-260-2
Perfect score: 2671
Sequence: 1 MAGRESLVSWMRAFHGCDSEAE.....PASVSGQADADSTSPRSQEP 507

Scoring table:
BLOSUM62:
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 segs, 1215238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/gcn2.1/USPTO_sppol/US9977260/runat_29072003_092741_14636/app.query.fasta.1.647
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=fto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09977260.ecgn.1.1.4237 -runat.29072003_092741_14636 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOS
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlinu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gp_estcl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_esthum:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*

29: gb_gss2:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2287.5	85.6	2046	11 AK013606	AK013606 Mus muscu
2	1978	74.1	2001	11 AK079816	AK079816 Mus muscu
3	1510.5	56.6	1043	12 BC054024	BC054024 AGENCOURT
4	1421	53.2	967	9 AL538511	AL538511 AL538511
5	1360.5	50.9	936	13 BU184285	BU184285 AGENCOURT
6	1330	49.8	1035	9 AL580639	AL580639 AL580639
7	1288.5	48.2	986	13 BX369553	BX369553 BX369553
8	1231	46.1	1040	9 AL558929	AL558929 AL558929
9	1226.5	45.9	1129	12 BM450510	BM450510 AGENCOURT
10	1216	45.5	1032	9 AL558805	AL558805 AL558805
11	1209	45.3	906	10 BG744770	BG744770 602722891
12	1207	45.2	1012	9 AL580543	AL580543 AL580543
13	1193.5	44.7	822	10 BC395801	BC395801 602458975
14	1156	43.3	949	9 AL567349	AL567349 AL567349
15	1139	42.6	837	12 BI912704	BI912704 603176356
16	1134	42.5	772	10 BG756878	BG756878 602713275
17	1127.5	42.2	708	12 BM715548	BM715548 UT-E-EIO
18	1119.5	41.9	805	10 BE781641	BE781641 601470504
19	1116	41.8	1138	10 BE791829	BE791829 601586106
20	1114	41.7	741	12 BM719266	BM719266 UT-E-EIO
21	1093	40.9	953	12 BI488576	BI488576 603021104
22	1087	40.7	931	12 BI549778	BI549778 603194514
23	1080	40.4	862	12 BI768288	BI768288 603056458
24	1078	40.4	607	10 BE257861	BE257861 601109450
25	1074	40.2	1102	9 BM561134	BM561134 AGENCOURT
26	1071	40.1	729	12 AV702542	AV702542 AV702542
27	1069	40.0	735	10 BE779216	BE779216 601464660
28	1064	39.8	1000	14 BE712568	BE712568 60112568
29	1060.5	39.7	950	10 BE795492	BE795492 601590164
30	1019	38.2	624	12 BI133187	BI133187 UT-M-BH3
31	997	37.3	631	10 BE254257	BE254257 601114671
32	995.5	37.3	1130	12 BM552996	BM552996 AGENCOURT
33	990.5	37.1	912	12 BI819535	BI819535 603036746
34	986.5	36.9	3064	11 AK037036	AK037036 Mus muscu
35	985	36.9	913	13 BQ954082	BQ954082 AGENCOURT
36	976	36.5	1171	12 BM918304	BM918304 AGENCOURT
37	973	36.4	952	10 BE791472	BE791472 601586006
38	958.5	35.9	1014	12 BM925586	BM925586 AGENCOURT
39	956	35.8	591	10 BE257454	BE257454 601111403
40	943	35.3	1075	11 AK018182	AK018182 Mus muscu
41	937	35.1	824	13 BU848605	BU848605 AGENCOURT
42	927	34.7	569	14 CD363397	CD363397 UT-M-GL0
43	927	34.7	800	10 BG703256	BG703256 602685347
44	920	34.4	1080	12 BM476098	BM476098 AGENCOURT
45	913	34.2	552	12 BM701130	BM701130 UT-E-EIO

ALIGNMENTS

RESULT 1
LOCUS AK013606
DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length
tyrosine kinase, clone:2900029A15 product:megakaryocyte-associated
tyrosine kinase, full insert sequence.
ACCESSION AK013606
VERSION AK013606.1 GI:12851034
KEYWORDS HTG; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL MEDLINE
PUBMED 99279253
REFERENCE 10349636
AUTHORS

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE
PUBMED 20499374
REFERENCE 11042159
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishie, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Motani, K.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multiplexed sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE
PUBMED 20530913
REFERENCE 11076861
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
Saito, T., Okazaki, Y., Gojodori, T., Bono, H., Kasukawa, T., Saito, R.,
Kado, K., Matsuda, H., Ashburner, M., Batalov, S., Casavola, T.,
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nakado, I., Pesole, G.,
Quackenbush, J., Schiraldi, L. M., Stuhli, F., Suzuki, R., Tomita, M.,
Wagner, T., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hotman, M., Hume, D. A., Kamlay, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-Oka, K., Wang, K. H., Weitz, C., Wittaker, C., Wilmink, L.,
Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

JOURNAL MEDLINE
PUBMED 21085660
REFERENCE 11217851
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED 12187851
REFERENCE 11217851
AUTHORS

TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
JOURNAL

[illegible]

Db 208 CCCGGTAGCTAACCTTTCGAAAGGCTGACATGGTGCACCATCTTGAGGCGCTGTGAGGAC 267
 Qy 84 LysSerTrpYrArgValLysHisHisThrSerGlyngluGlyLeuLeuAlaGly 103
 Db 268 AAGAGCTGTACGACGACCAACACCATGGCGAGGAGGAGGCTGTGCGGCGCT 327
 Qy 104 AlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetProTrpPhe 123
 Db 328 GCTGTGGACAGACGGGAGGCGCTCTCCACAGACCCCAAGCTCAGCTCATCTCATGTGTTT 387
 Qy 124 HisGlyLysIleSerGlyngluAlaValGlnGlnLeuGlnProProGluAspGlyLeu 143
 Db 388 CATGGCAAGATCTCCGGCCAGGAAGACATACAGCAGCTGACGCCACCGGAGAGGCGTG 447
 Qy 144 PheLeuValArgGluSerAlaArgHisProGlyAspTrpYrValLeuGlyValSerPheGly 163
 Db 448 TTCCTGTGAGGGAATCAGCTGCTCACCCTGGAGACATCTGTGTGTGTGTGTGTGCGC 507
 Qy 164 ArgAspValIleHisTrpYrArgValLeuHisArgAspGlyHisLeuThrIleAspGluAla 183
 Db 508 CGTGACGTCATCCACTACCGCTGTGGCATGAGATGGGACACTCAGCATGATGAGGCC 567
 Qy 184 ValPhePheCysAsnLeuMetAspMetValGlnHisTrpYrSerLysAspGlyAlaIle 203
 Db 568 GTGTGTTCTGTAACTGTATGACATGGTGGAGCAGTACACCAAGGACAAAGGGCGCATC 627
 Qy 204 CysThrLysLeuValArgProLysArgGlyHisGlyThrLysSerAlaGluGluLeu 223
 Db 628 TGCACCAAGCTGTGTGAAGCCAGAGGACGAAAGGCGCAAGTGTGCAGAGGAGAGCTC 687
 Qy 224 AlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGlu 243
 Db 688 GCCAAGCTGTGCTGCTACTGCTGACATGCTGACATGCTGGAGACAGATTTGGAGAG 747
 Qy 244 GlyGluPheGlyAlaValLeuGlnGlyLysTrpLeuGlyGlnLysValAlaValLysAsn 263
 Db 748 GGGGAGTTTGGAGCGCTCTCAACAGGGTGAGTACCTGGAGCAGGAGGCGCTGAGAGAT 807
 Qy 264 IleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMetThrLysMet 283
 Db 808 ATCAAGTGTGATGTGACAGCCGACGCTCTCTGATGAGAGCGCTGTATACCAAGCTG 867
 Qy 284 GlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeuTrpIleVal 303
 Db 868 CAGCAGACAGACCTAGTCCGACTCTGGGTGTGATCTGCACCAAGGCTGTACATGTCTC 927
 Qy 304 MetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGlyArgAlaLeu 323
 Db 928 ATGAGCAGCTGTGACCAAGGCAACCTGTGTAACTCTGCGCACGCGGGCGCTGCTCTT 987
 Qy 324 ValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMetGluTrpLeu 343
 Db 988 GTGAGCAGCTCTCAGCTTCTTCTGATTTGCTTCTGTAAGGATGGAATACCTG 1047
 Qy 344 GluSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSerGluAsp 363
 Db 1048 GAGAGCAAGAGCTGTGTCCACCGGAGCTGCTGCTGCGAGACATCTGCTGTGAGGAC 1107
 Qy 364 LeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGlyLeuAspSer 383
 Db 1108 TTGGTGGCCAAAGGTGAGTGTGGCTTGTACCAAGGAGAGGCCCAAGGGCTGTGAGTCA 1167
 Qy 384 SerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLysPheThrSer 403
 Db 1168 AGCGGCTGTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1227
 Qy 404 LysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrpGlyArgAla 423
 Db 1228 AAGTCGATGCTGTGAGTTTGGGGTGTGTGTGGGAAAGTCTCTTATGAGAAAGACC 1287
 Qy 424 ProTrpProLysMetSerLeuLysGluValSerGluAlaValGluLysGlyTrpYrArgMet 443
 Db 1288 CCATACCCCAAGATGTGCTTAAAGAGGTTCAGAGGCTGTGAGAGAGGTTTACCGCATG 1347

Qy 444 GluProProGluGlyCysProGlyProValHisValLeuMetSerSerCysTrpGluAla 463
 Db 1348 GAGCCCCCATGAGTGGCTCCAGGCTGTGTGACACACCTCATGTGCTACTGCTGGAGGCA 1407
 Qy 464 GluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArgGluLeuArg 483
 Db 1408 GAGCTGTGCGGCGCCAGCACCTTCCGCAAAATAGTGAAGGAGGAGGCGGAGACTCCGC 1467
 Qy 484 SerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThrSerProArg 503
 Db 1468 AGTGTGGTGTCTTCGCGCCCGCTGTGGGAGACAGGAGGCTGAGGCTGACCTCCACACGG 1527
 Qy 504 SerGlnGluPro 507
 Db 1528 AGCCAGGAGACCC 1539
 RESULT 2
 AK079816
 LOCUS
 DEFINITION
 Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:4430076C03 product:megakaryocyte-associated tyrosine kinase, full insert sequence.
 ACCESSION
 AK079816
 VERSION
 AK079816.1 GI:26098774
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL
 MEDLINE
 PUBMED
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, N., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoke, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, O., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer
 JOURNAL
 MEDLINE
 PUBMED
 20530913
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaoka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schiml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hotmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Rling, B., Rlingwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

FEATURES	Source
BASE COUNT	193 a 297 c 330 g 147 t
ORIGIN	
COMMENT	On Feb 13, 2001 this sequence version replaced gi:12802004.
JOURNAL	
TITLE	Full-length cDNA libraries and normalization
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
REFERENCE	1. (bases 1 to 967)
ORIGIN	
KEYWORDS	EST.
VERSION	AL538511.2 GI:31263101
ACCESSION	AL538511
DEFINITION	AL538511 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
LOCUS	AL538511 967 bp mRNA linear EST 31-MAY-2003
RESULT 4	
Db	903 TGGAAATCCCTTGAAAGACCAAAAGCTTTGGACCCCAACCTGGCCGCC 953
Oy	340 etglu-tyrleuglser--lyslysleuvalhisargasphealaala 355
Db	843 CGAGCCCTCGTGAACACCGCTCANCTCTGACATTTCTCTGCACCTGGCCCAAGCA 902
Oy	321 Argala-LeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAla-leuglylm 340
Db	783 CATTGTCTGTGAGCGACCTGACCAAGGGCAACCTGGGTGAACCTTCTTGGGACCCGGGGGT 842
Oy	301 rllvalmetcunhisvalserlysglyasnleu-ValAsnpheluarghrarg-gly 320
Db	723 AAGATGCACAACGAGAACCTGTGGCGCTCTCTGGGCGGTGATTCCTGCACCAAGGGCTGTA 782
Oy	282 Lysmetcunhisglu-AsnleuValArgleuLeuGlyValIleleuHisGlnGlyLeuTy 301
Db	663 AAGATATCAAGTGTGATGTACACGCCAGGCTCTTCGGACGAGACGCCGCTCATGACG 722
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	262 LysAsnIlelyscysAspAspValThrAlaGlnIlePheLeuaspGluThrAlaValmethr 281
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATCCTGGGGCAAAAGTGGCCGTG 662
Oy	242 GlygluGlyIupheGlyAlaValleuGlnGlyIuTyrrleuGlyGlnLysValAlaVal 261
Db	603 GGAGGGGAGAGTTTGGAGCTGTCTCTCGAGGGTGAATC

Alignment Scores:	1,78e-120	Length: 967
Pred. No.:	1421.00	270
Score:	94.10%	Conservative: 1
Percent Similarity:	93.75%	Mismatches: 5
Best Local Similarity:	53.20%	Indels: 12
Query Match:	9	Gaps: 1
DB:		
US-09-977-260-2 (1-507) x AL538511 (1-967)		
QY	2 ALaGlyARgSLySerLeuValSerTPRarGaLapheNlSGLyAspSerAlaGlu 21	
Db	139 GCGGGGGCGGCGC-----TGGGGGGCGGCC 167	
QY	22 LeuProARgYalSerProARgPheLeuAlaATPRhISProProProValSerAlaArg 41	
Db	163 CCGAGAGAGGTAGAGCCCGGCTTCCTCCAGGCGTGGACACCCCGCTCCGTCAGAGG 222	
QY	42 MetProthARgARgTPRAlaProGlyThrGIncyLIethrLyASgAlNHISThrArg 61	
Db	223 ATGCAACAGAGGGCGTGGGGCCCGGGACCCAGTGTATCACCAATGGAGACACCCGC 287	
QY	62 ProLysProGlyGluLeuAlaPheARgLySGLyAspValThrILeUGluAlaCys 81	
Db	283 CCCAAGCCAGGGAGAGCTGGCTCCGCAAGGGCGAGCTGGTACCATCTCGAAGGCTGCG 347	
QY	82 GluAsnLysSerTPRrPrARgYalLysHISThrSerGInGInGlyLeuLeuAla 100	
Db	343 GAGAAACAAGAGCTGGTACCGCGTCAAGCACACACAGTGGAGGAGAGGGCTGTGGCA 407	
QY	102 AlaGlyAlaLeuARgLyGluARgAlaLeuSerAlaAspProLysLeuSerLeuMetPro 121	
Db	403 GCTGGGGCGGTGGGGGAGGGGAGGGCGCTCTCGGAGACCCCAAGCTAGGCTCATGGCG 466	
QY	122 TrpPheHISGLyLysILeserGlyGInGluAlaValGInGInLeuGInProProGluAsp 141	
Db	463 TGGTTCACACGGAGAAAGTCTCGGGGCGCAGAGGCTGTCCAGAGCTGCACCTCCGAGAGAT 522	
QY	142 GlyLeuPheLeuValARgLySerAlaARgHISProGlyAspTyValLeuCysValSer 161	
Db	523 GGGCTGTCTCTGGTGGGGAGTCCGGCGGCCACCCGGCGACTACGTCCTGTGGGTAGGC 582	
QY	162 PheGlyARgAspValILeHISTyTrARgValLeuHISARgAspGlyNHISLeuThrILeAsp 181	
Db	583 TTGGCGCGGAGCTATCCACTACCCCGTGGCTGCACCCGAGCGCACTACAAATGGAT 642	
QY	182 GluAlaValAPhePheCysAsnLeuMetAspMetValGInHISTySerLyAspLysGly 201	
Db	643 GAGGCGGTGTTCTTCGCAACCTCATGACATGAGTGGAGACTTACAGCAAGACAAAGGCG 702	
QY	202 AlaILeCysThrLyLeuValARgProLysARgLySHISGlyThrLysSerAlaGlu 221	
Db	703 GCTATCTGCACCAAGTGTGAGACCAAGAGCGGAACCGGAGCCCAACTCGCGGAGAG 762	
QY	222 GluLeuAlaARgAlaGlyTrpLeuLeuAsnLeuGInHISLeuThrLeuGlyAlaGlnILe 241	
Db	763 GAGCTGGCCAGGGCGGGCGTGTACTGAACCTGCAGCATTTGACATTGGAGGACAAATC 822	
QY	242 GlyGluGlyLeuPheGlyAlaValLeuGInGlyLutTyLeuGInGlyLysValAlaVal 261	
Db	823 GAGAGAGGAGAGTTTGGAGCTCTCTCGCAGGGTGAAGTACGTGGGGCAAAAGTGGCGGTG 882	
QY	262 LysAsnILeLysCysAspValThrARgAlaGlnAPheLeuAspGlyThrAlaValMetThr 281	
Db	883 AAGAAATATCAAGTGTGATGTGACACCCAGGCGCTTCCrGAGAGAGAGCGCGTGTATGACG 942	
QY	282 LysMetGlnHISGluAsnLeuVal 289	
Db	943 AAGATGCACACGAGAACTGGTGG 966	

LOCUS BU184285 936 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7942749 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6171643
5', mRNA sequence.
ACCESSION BU184285
VERSION BU184285.1 GI:22698269
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 936)
NIH-MGC <http://mhc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM13540 row: e column: 20
High quality sequence stop: 674.
Location/Qualifiers
1..936
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6171643"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: PCMV-SPO6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo
Average insert size 1.75 kb. library constructed by Life
Technologies."
BASE COUNT 184 a 268 c 314 g 169 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 6,35e-115 Length: 936
Score: 1360.50 Matches: 288
Percent Similarity: 77.17% Conservative: 6
Best Local Similarity: 75.59% Mismatches: 11
Query Match: 50.94% Indels: 77
DB: 13 Gaps: 3
US-09-977-260-2 (1-507) x BU184285 (1-936)
OY 99 LeuLeuAlaAlaGlyAlaLeuAlaGluAlaLeuSerAlaAspProLysLeuSer 118
DB 11 GTGCTGACAGCTGGGGGCTGCGGAGCGGAGGCCCTCTCCGACGCCAAGCTCAGC 70
OY 119 LeuMetProTrrPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnPro 138
DB 71 CTCATGCCGTGCTCCACGGGAATCTCGGGCCACGAGCTGTCCACACTGACGCT 130
OY 139 ProGluAspGlyLeuPheLeuValAlaGluSerAlaArgHisProGlyAspTyrValLeu 158
DB 131 CCCGAGATGGGCTGTCTCTGTCGCGAGATCCGCGGCCACCCCGGCGACTACGCTCTG 190
OY 159 CysValSerPheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeu 178
DB 191 TGCCTGACCTTGGCCCGACGATCATCACCGCGGTGACACCGGAGGCGCACCTC 250
OY 179 ThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrSerLys 198
DB 251 ACATCATGAGGCGCGTCTTCTGCAACTCATGACATGATGAGAGCATTTACAGCAG 310
OY 199 AspLysGlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSer 218

DB 311 GACAAGGCGCTATCTGCACCAAGCTGTGTGACCAAAAGCGGAACAGGAGCAAGTCG 370
OY 219 AlaGluGluGluLeuAlaAlaArgAlaGlyTrrPheLeuAsnLeuGlnHisLeuThrLeuGly 238
DB 371 GCCGAGGAGGAGCGGCGGAGGCGGCGGTGTACTGAAACCTGCAGCATTTACATTGGGA 430
OY 239 AlaGlnIleGlyGlnGlyGluPheGlyAlaValLeuGlnGlnGlyLysGlyGlnLys 258
DB 431 GCACGATCGGAGAGGAGAGATTGTGAGCTCTCTGACGGGTGAGTACTGTGGGCAAAAG 490
OY 259 ValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGlyThrAla 278
DB 491 GTGGCGCGAAGATATCAAGTGTGATGTGACACCCGAGGCTCTCTGACGAGAGCGCC 550
OY 279 ValMetThrLysMetGlnHisGlnAsnLeuValArgLeuLeuGlyValIleLeuHisGln 298
DB 551 GTCATGACGAAGATGACCAACGAAACCTGTGGCTCTCCGCGGCGGTATCTCCACACG 610
OY 299 GlyLeuTrrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThr 318
DB 611 GGGCTGTACATTTGTATGAGAC-----GGCAACCTGTGMACTTTCTCGGAGC 661
OY 319 ArgGlyArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlu 338
DB 662 CGGGGTGAGCCCTCGTGAACACCGCTCAGCTCTGCGAGTTTCTCT----- 708
OY 339 GlyMetGlyTrrLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIle 358
DB 708 ----- 708
OY 359 LeuValSerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArg 378
DB 708 ----- 708
OY 379 LysGlyLeuAspSerSerArgLeuProValLysTrrPheAlaProGluAlaLeuLysHis 398
DB 708 ----- 708
OY 399 GlyLysPheThrSerLysSerAspValTrrPheSerGlyValLeuLeuTrrGluVal 417
DB 709 ---AAGTTCACACCAAGTCTGATGCTGGAGGTTTGGGGGCTGTGGGAGGTC 765
OY 418 PheSerTyrGlyArgAlaProTrrProLysMetSerLeuLysGluValSerGluAlaVala 437
DB 766 TTCTCATTTGACACCGGCTCCGTACCTTAANAATGCTCTGAAGAAGGTGTGAGAGCCCT 825
OY 437 GlnLysGlyTyrArgMetGluProGlnGly---CysProGlyProValHisValLe 456
DB 826 GGAAGAAGGCTACCGCATGGAACCCCGGAGGCTGTGACGGGCCCGGCGACAGCTCT 885
OY 456 Umet-SerSerCysTrrGluAlaGlu---ProAlaArgArgProPhe 471
DB 886 CATGAACAACTGCTGGGAACAAACCCCGGCGGCGGCGGCTTT 934
RESULT 6
LOCUS AL580639/c 1035 bp mRNA linear EST 01-JUN-2003
DEFINITION AL580639 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ007K05 3-PRIME, mRNA sequence.
ACCESSION AL580639
VERSION AL580639.2 GI:31318906
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1035)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12946853.
Contact: Genoscope
Genoscope - Centre National de Sequencage

Best Local Similarity: 86.75% Mismatches: 20
 Query Match: 48.24% Indels: 19
 DB: 13 Gaps: 3
 US-09-977-260-2 (1-507) x BX369553 (1-986)

QY 195 HistSerLysAsp-LysGlyAlaIleCysThrLysLeuValArpProLysArgLysH1 214
 DB 3 CATTACGCAAGGCAAGGCGCTATCTGCACCACTGCTGTGACCAAGCGGAACA 62
 QY 214 sGlyThrLysSerAlaGluGluLeuAlaArgAlaGlyTrpLeuLysLeuGlnH1 234
 DB 63 CGGACCAAGTGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122
 QY 234 sLeuThrLeuGlyAlaGlnIleGlyGluGluGluGluGluGluGluGluGluGlu 254
 DB 123 TTTCACATTTGGACACAGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 182
 QY 254 rLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnIle 274
 DB 183 CTGGGGCAAAAGGTGGCCGCGAGATATCAAGTGTGATGTGACAGCCAGGCTTCT 242
 QY 274 uAspGluThrAlaValMetThrLysMetGlnHisGlnAsnLeuValArgLeuGly 294
 DB 243 GGACGACGCGCCCTCATGACGAGAGATGCAACACAGAGAGAGAGAGAGAGAG 302
 QY 294 lIleLeuHisGlnGlyLeuTrpIleValMetGlnHisValSerLysGlnLysLeuVal 314
 DB 303 GATCTCACAGGAGGCTGTACATGTCATGAGACAGTGTGAGAGAGAGAGAGAGAG 362
 QY 314 nPheLeuArgThrArgLysArgAlaValAsnThrAlaGlnLeuGlnPheSerLe 334
 DB 363 CTTCCTCGGAGCCGGGGGTGAGCCCTGCGTAACAGCGCTAGTCTCGAGCTTCT 422
 QY 334 uHisValAlaGluGlyMetGluTrpLeuGlnSerLysLysLeuValHisArgAspLeu 354
 DB 423 GCACGTGGCGAGGCGCATGAGTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
 QY 354 aAlaArgAsnIleLeuValSerLysValAlaValLysValSerAspPheGlyLeu 374
 DB 483 CGCCCGCAACATCTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 542
 QY 374 aLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThrAlaPro 394
 DB 543 CAAGCGGAGGAGAGAGGCGTACAGCTCAGCGGCTGCGTCAAGTGAAGGCGCGA 602
 QY 394 uAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSer-PheGlyValLeu 414
 DB 603 GGCTCTCAACACGCGGAGATTACACAGCAAGTGTGAGATTGTTGGGTGCTGC 662
 QY 414 eUTrPGluValPheSerTrp---GlyArgAlaProTrpProLysMetSerLeuLysGlu 433
 DB 663 TCTGGGAGGCTCTTCTCATATGGAGGAGGCTCGTCCGTAATGTCATGTAAGAGG 722
 QY 433 aLysGlnAlaValGluLys-GlyTrpArgMetGluProProGluGly-CysProGlyP 452
 DB 723 TGTGGAGGCGCCGCGAGATAGGTTACCGCATGGAACCCCGAGGCTGTGCAAGCC 782
 QY 452 rOValHisValLeu-MetSerSer-CysTrpGluAlaGlu--ProAlaArgArgProPro 470
 DB 783 CCGGACACCTCTTCATGAGAGAGCTTGTGGAGAGAGAGAGAGAGAGAGAGAGAG 842
 QY 471 PheArgLysLeu-AlaGluLysLeuAla-ArgGluLeu-ArgSerAlaGlyAlaProAla 489
 DB 843 TTTCGGAAGTGTGCGAGAACTTGTCCGAGAGCTTTCGCAATGGGCGGTGGCCAG 902
 QY 490 Ser-----ValSerGlyGlnAspAlaAspLys-----SerThrSerProArg 503
 DB 903 TCCCTTGGTTTAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 962
 QY 504 SerGlnGluPro 507
 DB 963 AGTCCTGTCC 974

RESULT 8
 AL558929
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1040)
 L.W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 15, 2001 this sequence version replaced gi:12903930.
 COMMENT
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqrel@genoscope.cns.fr, Web : www.genoscope.cns.fr
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 9238.r For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSDDJ007AF03QPI&cluster=9238.r. Contact :
 Feng Liang Email : fliang@life.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSDDJ007AF03QPI.
 Location/Qualifiers
 1..1040
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSDDJ007YK05"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

BASE COUNT 188 a 341 c 338 g 162 t 11 others
 ORIGIN

Alignment Scores:
 Pred. NO.: 6.25e-103 Length: 1040
 Score: 1231.00 Matches: 231
 Percent Similarity: 98.72% Conservative: 1
 Best Local Similarity: 98.30% Mismatches: 3
 Query Match: 46.09% Indels: 1
 Gaps: 0

US-09-977-260-2 (1-507) x AL558929 (1-1040)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 DB 335 ATGGCGGGGCGAGGCTCTGTGGTTCTCGGGGCGCATTTTCAGGCTGTGATTCGCTGAG 394
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla 40
 DB 395 GAACCTCCCGGGGAGAGCCCGCTTCTCGAGCTGTGACACCCCTCCGCTGTGAGCC 454
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnHisThr 60
 DB 455 AGAATGCCAACGAGCGGTGGGCGCGGAGCAACAGTGTATCACAATGCGACACACC 514
 QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGlnAla 80
 DB 515 CGCCCAAGCGAGGAGGAGTGGCTTCGCAAGGAGCGAGTGTACACATCTGTGAGGCC 574

QY 81 CysGluAsnLysSerTrpTyrArgValLysHisThrSerGlyGlnGlyLeuLeu 100
|||||
Db 575 TGCAGAACAGAGCTGTACCGCTCAGACACACCACTGACAGGAGGCTGCTG 634
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
|||||
Db 635 GCACCTGGGGGGCTGGCGGAGCGGAGCCCTCTCCCGACAGCCCAAGCTCAGCTCATG 694
QY 121 ProTrpPheHisGlyLysLysSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
|||||
Db 695 CCGTGGTTCACAGGAGATCTCGGCGAGAGGCTGTCCAGCAGCTGACGCTCCCGAG 754
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuCysVal 160
|||||
Db 755 GATGGCTGTCTTCGCTGCGGAGCTCCGCGCGACCCCGGCGACTACGCTCTGTGCTG 814
QY 161 SerPheGlyArgAspValLysHisTyrArgValLeuHisArgAspGlyHisLeuThrIle 180
:::
Db 815 ACCCTT-GGCGGCSAGCTATCCACTACCGCTGCTGCACCGCGACGCTCAGCTCAGATC 873
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys 200
|||||
Db 874 GATGAGCGCTGTCTCTCTCCTCAACCTCATGTGATGTCGATTCACAGCAAGGACAG 933
QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
|||||
Db 934 GCGGCTATCTGCACCAAGCTGTGTAGACCAAGCGAAGACGCGGACCAAGTGGCGGAG 993
QY 221 GluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGlnHisLeu 235
|||||
Db 994 GAGGAGCTGGCAGGCGGCTGTACTGACCTGACGATTTG 1038

RESULT 9
BM450510 1129 bp mRNA .linear EST 05-FEB-2002
DEFINITION AGENCOURT_6394260 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5494913
5', mRNA sequence.
BM450510
VERSION BM450510.1 GI:18499550
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1129)
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12120 row: p column: 18
High quality sequence stop: 600.
Location/Qualifiers
1..1129
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5494913"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

FEATURES
source
BASE COUNT 233 a 345 c 369 g 182 t

ORIGIN
Alignment Scores:
Pred. No.: 1,82e-102 Length: 1129
Score: 1226.50 Matches: 271
Percent Similarity: 87.46% Conservative: 8
Best Local Similarity: 84.95% Mismatches: 26
Query Match: 45.92% Indels: 14
DB: 12 Gaps: 8
US-09-977-260-2 (1-507) x BM450510 (1-1129)
QY 89 ValLysHisThrSerGlyGlnGlyLeuAlaAlaGlyAlaArgGlu 108
Db 6 GTCAAGCACACACCACTGTGACAGAGGAGGCTGTGACAGTGGCGGAGCGG 65
QY 109 GluAlaLeuSerAlaAspProLysLeuSerLeuMetProTrpPheHisGlyLysLys 128
Db 66 GAGGCCCTCTCCGACAGACCCCAAGCTCAGCTCATGCGGTGTTCCACGGGAAGATCTCG 125
QY 129 GlyGlnGluAlaValGlnGlnLeuGlnProProGluAspGlyLeuPheLeuValArgGlu 148
Db 126 GCGCAGAGGCTGTCCAGACGCTGCACCTCCGAGAGATGGCTGCTCCGTGCGGAG 165
QY 149 SerAlaArgHisProGlyAspTyrValLeuCysValSerPheGlyArgAspValLysHis 168
Db 186 TCCGCGGCCACCCCGCGACTACGCTGTGCTGCTGAGCTTGGCCGCGACGTCATCCAC 245
QY 169 TyrArgValLeuHisArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsn 188
Db 246 TACCGCTGTGACCGCGCGAGCGGCTCAGCAATCATGATGAGCGCTTCTCTGCAAC 305
QY 189 LeuMetAspMetValGluHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuVal 208
Db 306 CTCATGACATGTGTGAGCATTTACAGCAAGACAGAGCGGCTATCTGCACCAAGCTGTG 365
QY 209 ArgProLysArgLysHisGlyThrLysSerAlaGluGluLeuAlaArgAlaGlyTyr 228
Db 366 AGACCAAGCGGAAACAGCGGACCAAGTCCGCGAGAGAGCTGCGCAGCGGCGCTGG 425
QY 229 LeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGlnGlyPheGlyAla 248
Db 426 TTACTGACCTGACGACATTTGACATTTGGAGCACAGTCCGAGAGGAGGATTGGAGCT 485
QY 249 ValLeuGlnGlyLysTyrLeuGlnGlyGlnLysValAlaValLysAsnIleLysCysAspVal 268
Db 486 GTCTCAGAGGTGATGACCTGGGCGGAAAGGTGGCGGTGAAGATTCAGATGTGATGTG 545
QY 269 ThrAlaGluAlaPheLeuAspGluThrAlaValMetThrLysMetGlnHisGluAsnLeu 288
Db 546 ACAGCCGAGGCTTCTGTGACGAGCGCGCTCATGACGAAGATGCACACGGAACCTG 605
QY 289 ValArgLeuLeuGlyValLysLeuHisGln-GlyLeuTyrIleValMetGlnHisValSer 308
Db 606 GTCCGCTCTCTGGCGCTGATCTCCGACACAGGGGCTGTACATTTGTATGAGCACCTGAG 665
QY 308 rLysGlyAsnLeuValAsnPheLeuArg-ThrArgGlyArgAla-LeuValAsnThrAla 327
Db 666 CAAGGCAACCTGTGTGATCTTCTGCGGAGACCGGGGCTGAGGCGCTCGGAACACGCTG 725
QY 328 Gln-LeuLeuGln---PheSerLeuHisValAlaGlnGly---MetGlyTyrLeu---G1 344
Db 726 CAGCTTCCTGAGGATTTCTGTGCGCCGCGGAGGAGGATTTGAGATTTACCTGCGGAGA 785
QY 344 uSerLysLysLeuValHisArg-AspLeuAla-----AlaArgAsnIleLeuValSerG 362
Db 786 ACAAGAACTTTGTACACCGCGGACCTGGCCCGCGCAATTCCTGCGCCCAAG 845
QY 362 LysPheLeuAlaLysValSer---AspPheGlyLeuAlaLysAlaGluArg---LysG 380
Db 846 GACCTGGGGGCGCAAGTGCAGCGGACCTTGGGCGCTGCGCAAAAGCCCAAGCGAGG 905
QY 380 LysLeuAsp---SerSerArgLeu---ProValLysTrpThrAlaPro 393

into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC library."

BASE COUNT 142 a 306 c 281 g 176 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 5,47e-101 Length: 906
Score: 1209.00 Matches: 241
Percent Similarity: 97.98% Conservative: 1
Best Local Similarity: 97.57% Mismatches: 5
Query Match: 45.26% Indels: 1
DB: 10 Gaps: 0

US-09-977-260-2 (1-507) x BG744770 (1-906)

OY 261 VALLYSANLLELYSCYSPVALTHRALAGLALPHELEUNSPGLTHRALAVALMET 280
DB 899 GTTAAAGATATCAAGGCTGTGG-ACAGCCAGGCTTCCCTGGAGAGAGCCCGTCATG 841
OY 281 ThrlyMetGlnHisGlnLeuValArgLeuGlyValIleLeuHisGlnGlyLeu 300
DB 840 ACCAAGATGCACAGAGAACCTGGTGGCTCTCGGCGTGATCTGCACCGAGGAGCTG 781
OY 301 TTTTleValMetGlnHisValSerLysGlyAsnLeuValAsnPheluarqThrrArgly 320
DB 780 TACATGTCTATGAGACGACGACGACGACGACGACGACGACGACGACGACGACG 721
OY 321 ArgAlaLeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGlnGlyMet 340
DB 720 CGAGCCCTGTGTAACCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 661
OY 341 GlnTyrLeuGlnSerLysLeuValHisArgAspLeuAlaIleArgAsnIleLeuVal 360
DB 660 GAGTACCTGGAGAGCAGAGAGCTTGTGCACCGGACCTGGCCGCCGACACATCTGTGTC 601
OY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGlnArglyGly 380
DB 600 TCAGAGACCTGTGGCCCAAGGTCAGCAGCTTTGGCGTGGCCCAAGCCGAGCGGAGGG 541
OY 381 LeuAspSerSerArgLeuProValLysTrrPthAlaProGlnAlaLeuLysHisGlyLys 400
DB 540 CTAGACTCAAGCCGCTGCCCTCAAGTGAAGCGCCGACGCTCTCAACACGCGGAG 481
OY 401 PheThrSerLysSerAspValTrrPserPheGlyValIleLeuTrrPgluValPheSerTyr 420
DB 480 TTCACACGACAGTGGATGTCTGGACTTTTGGGGTCTGCTGCGAGAGGTCTTCTCATAT 421
OY 421 GlnArgAlaProTyrProLysMetSerLeuLysGlnValSerGlnAlaValGlnLysGly 440
DB 420 GGAGGGGCTCCGTAACCTTAATGTCATGAAAGAGGTGTGGAGGCCGCGGAGAGGG 361
OY 441 TyrArgMetGlnProProGlnGlyCysProGlyProValHisValLeuMetSerSerCys 460
DB 360 TACCGATGGAAACCCCGCAGGGCTGTCCAGGCGCCGCTGACCTCTCATGACACCTGC 301
OY 461 TrrPgluAlaGlnProAlaArgArgProPheArgLysLeuAlaGlnLysLeuAlaArg 480
DB 300 TGGAGAGCGAGAGCCCGCCGCGGACACCTTCCGAAACTGGCCAGAAAGCTGGCCCG 241
OY 481 GlnLeuArgSerAlaGlnAlaProAlaSerValSerGlyGlnAspAlaSpGlySerThr 500
DB 240 GAGCTACGACAGTGCAGAGTGCACCGCTCCGCTCTCAAGGAGAGAGCCGACGCTCAC 181
OY 501 SerProArgSerGlnGlnPro 507
DB 180 TCGCCCGGAAAGCCAGAGCC 160

RESULT 12
AL580543/c

LOCUS AL580543 1012 bp mRNA linear EST 01-JUN-2003
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION Homo sapiens cDNA clone CS0DJ015Y001 3-PRIME, mRNA sequence.
VERSION AL580543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1012)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 16, 2001 this sequence version replaced g1:1294667.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9238.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015AH01NP1&cluster=9238.r. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ015AH01NP1.
location/Qualifiers
1..1012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015Y001"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

BASE COUNT 158 a 346 c 277 g 194 t 37 others
ORIGIN

Alignment Scores:

Pred. No.: 9.77e-101 Length: 1012
Score: 1207.00 Matches: 253
Percent Similarity: 93.77% Conservative: 3
Best Local Similarity: 92.67% Mismatches: 15
Query Match: 45.19% Indels: 6
DB: 9 Gaps: 1

US-09-977-260-2 (1-507) x AL580543 (1-1012)

OY 235 LeuThrLeuGlnGlyAlaGlnIleGlyGlnGlnPheGlyAlaValLeuGlnGlyTyr 254
DB 1006 ATTACATTGGAGACAGA---TCGAGAGAGGAGATTGGACT-GTCTGCA-GGTGAGTAC 952
OY 255 LeuGlnGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeu 274
DB 951 CT-GGGCAAAAGGTGGCCCGTGAAGAAATATCAATGTGATGTGACACCCAGGCTTCTCTG 893
OY 275 AspGluThrAlaValMetThrLysMetGlnHisGlnAsnLeuValArgLeuGlnGlyVal 294
DB 892 GACGACAGCGCGCTATGAGAGAGCAACAGCAACAACTGGTGGTGGTGGCGGCG 833
OY 295 IleLeuHisGlnGlyLeuTyrIleValMetGlnHisValSerLysGlyAsnLeuValAsn 314
DB 832 ATCTCTACACAGGGGCTGTATCATTCATGAGGACAGTGAAGAGGCAACCTGATGAC 773
OY 315 PheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeu 334
DB 772 TTTCTGGAGAGCCGGGGTGGAGCCCTGTAACACCGCTCAGCTCTGAGATTCTCTCG 713

/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH-MGC Library."

BASE COUNT 157 a 267 c 291 g 122 t
ORIGIN

Alignment Scores:
Pred. No.: 1.37e-94 Length: 837
Score: 1139.00 Matches: 230
Percent Similarity: 90.31% Conservative: 3
Best Local Similarity: 89.15% Mismatches: 12
Query Match: 42.64% Indels: 16
DB: 12 Gaps: 1

US-09-977-260-2 (1-507) x B1912704 (1-837)

OY 2 AlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 21
DB 88 GCGGGGCGCGGC-----TCGGGGCGCGCC 111
OY 22 LeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAlaArg 41
DB 112 CCGAGCAGGTGAGCCCTCTCCGAGCCTGGACCCCTCCCTCCGTCACGCCAGG 171
OY 42 MetProThrArgArgTrpAlaProGlyThrGlnCysIleThrIysCysGlnHisThrArg 61
DB 172 ATGCCAAGAGAGCGGTGGCCCGCGGACCCAGTGTATCACCMAATGGAGACACCCGC 231
OY 62 ProlYsProGlyGluLeuAlaPheArgIysGlyAspValValThrIleLeuGlnAlaCys 81
DB 232 CCCAAGCCAGGGAG-CTGGCCTCCGCAAGGGCGCGTGCACCATCTCGAGGCTGC 290
OY 82 GluAsnIysSerTrpTrpArgValIysHisThrSerGlyGlnGluLeuAla 101
DB 291 GAGAACAGAGCTGTACCGCTCAGCACCACCACTGAGAGGGGCTGTCGCA 350
OY 102 AlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProlYsLeuSerLeuMetPro 121
DB 351 GCTGGGGCGCTGCGGAGCGGAGGCCCTCTCCGAGACCCCAAGCTCAGCTCATGCCG 410
OY 122 TrpPheHisGlyIysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGluAsp 141
DB 411 TGGTCCACAGGAGATCTCGGCCAGAGGCTGCCAGAGCTGCAGCTCCGAGGAT 470
OY 142 GlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysValSer 161
DB 471 GGGGTGTCCTGTGGGAGTCCGGCGCCACCCGGGCACTACGTCTGTGCGTAGC 530
OY 162 PheGlyArgAspValIleHisTyrArgValLeuHisArgAspGlyHisLeuThrIleAsp 181
DB 531 TTGGCCGCGAGCTATCATCTACCGCTGCTGCACCGCAGCG-CACTTCACATCGAT 589
OY 182 GluAlaValIlePhePheCysAsnLeuMetAspMetValGluHisTyrSerIysAspIysGly 201
DB 590 GAGGCGGTGTCTCTGCACCTCATGTGAGCATTTACAGCAAGACAAATGGC 649
OY 202 AlaIleCysThrIysLeuValArgProIysArgIysHisGlyThrIysSerAlaGluGlu 221
DB 650 GCTATCTGCACCAAGCTGTAGACCAAGGAAACAGGGACCAAGTTCGGCCAGGAG 709
OY 222 GluLeuAlaArgAla-GlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnI 241
DB 710 GAGCTGGCCAGGGGGGCGCTGTTACTGACCTGAGCATTTGACATTGGAGACACAGAT 769
OY 241 eGlyGluGlyIlePheGlyAlaValLeuGlnGlyIleuTyrLeuGlyGlnIys 258
DB eGlyGluGlyIlePheGlyAlaValLeuGlnGlyIleuTyrLeuGlyGlnIys 258

DB 770 CGAGAGAGAGAGAGTTGGAGCTGC-CTGCAGGCTGAGTACTGGGGCAAAAGG 820

Search completed: August 1, 2003, 22:37:07
Job time : 3256 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 2, 2003, 04:55:27 : Search time 971 Seconds
(without alignments)
1077.180 Million cell updates/sec

Title: US-09-977-260-2
2671
Perfect score: 1 MAGRSLVSMRAPHGDSAE.....PASYSGDADGSTSPRQEP 507
Sequence:

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1439767 seqs, 1031500376 residues
Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2_1/USFPO/US09977260/runat_29072003_092742_14668/app.query.fasta_1.647
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=fnpb -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=D10sum62
-TRANS-human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pcp -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09977260.ecgn1.1.333 @runat_29072003_092742_14668
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUER -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2671	100.0	2000	9	US-09-977-269-1 Sequence 1, Appl1

2	2671	100.0	2000	10	US-09-977-260-1	Sequence 1, Appl1
3	2671	100.0	2000	11	US-09-977-261-1	Sequence 1, Appl1
4	2671	98.0	1989	14	US-10-084-817-341	Sequence 341, App
5	1247.5	46.7	2187	10	US-09-954-531-188	Sequence 188, App
6	1247.5	46.7	2187	15	US-10-298-377A-1	Sequence 1, Appl1
7	1247.5	46.7	2420	14	US-10-177-593-87	Sequence 87, Appl1
8	725.5	27.2	2015	10	US-09-954-456-1983	Sequence 1983, Ap
9	725.5	27.2	2015	14	US-10-007-010-3	Sequence 3, Appl1
10	715.5	26.8	1911	10	US-09-917-800A-1611	Sequence 1611, Ap
11	710	26.6	4517	15	US-10-298-377A-3	Sequence 3, Appl1
12	709.5	26.6	2298	14	US-10-175-523-50	Sequence 50, Appl1
13	701.5	26.3	5527	10	US-09-880-107-3710	Sequence 3710, Ap
14	681.5	25.5	2354	10	US-09-967-768A-300	Sequence 300, App
15	654.5	24.5	2770	9	US-09-977-269-5	Sequence 5, Appl1
16	654.5	24.5	2770	11	US-09-977-260-5	Sequence 5, Appl1
17	654.5	24.5	2770	11	US-09-977-261-5	Sequence 5, Appl1
18	654.5	24.5	2863	10	US-09-954-456-1631	Sequence 1631, Ap
19	654.5	24.5	7607	10	US-09-982-610-19	Sequence 19, Appl1
20	633.5	23.7	4221	14	US-10-220-801-8	Sequence 8, Appl1
21	630.5	23.6	6383	10	US-09-954-531-405	Sequence 405, App
22	614.5	23.0	2560	14	US-10-045-202-1	Sequence 1, Appl1
23	608.5	22.8	2582	14	US-10-220-801-4	Sequence 4, Appl1
24	608.5	22.8	2674	13	US-10-003-295-1	Sequence 1, Appl1
25	607.5	22.7	2451	10	US-09-771-161A-4	Sequence 4, Appl1
26	607	22.7	2468	14	US-10-045-202-3	Sequence 3, Appl1
27	587	22.0	2449	14	US-10-220-801-11	Sequence 11, Appl1
28	587	22.0	2456	13	US-10-186-399-1	Sequence 1, Appl1
29	587	22.0	2500	9	US-09-977-269-3	Sequence 3, Appl1
30	587	22.0	2500	10	US-09-977-260-3	Sequence 3, Appl1
31	587	22.0	2500	11	US-09-977-261-3	Sequence 3, Appl1
32	586.5	22.0	2520	11	US-09-953-047-3	Sequence 3, Appl1
33	582.5	21.8	3829	11	US-09-953-047-10	Sequence 10, Appl1
34	579.5	21.7	2604	14	US-10-171-581-317	Sequence 317, App
35	578.5	21.7	2627	9	US-09-962-436-313	Sequence 313, App
36	576.5	21.6	2184	10	US-09-728-952-82	Sequence 82, App
37	576.5	21.6	3454	13	US-10-044-900-48	Sequence 48, Appl1
38	576.5	21.6	3726	9	US-09-925-302-271	Sequence 271, App
39	563	21.1	3593	14	US-10-220-801-6	Sequence 6, Appl1
40	563	21.1	3663	10	US-09-919-172-84	Sequence 84, Appl1
41	558.5	20.9	3914	13	US-10-044-090-148	Sequence 148, App
42	556	20.8	2574	9	US-09-735-103-2	Sequence 2, Appl1
43	556	20.8	2574	13	US-10-045-428A-2	Sequence 2, Appl1
44	555	20.8	3306	11	US-09-954-556-10	Sequence 10, Appl1
45	551.5	20.6	2433	14	US-10-037-270-830	Sequence 830, App

ALIGNMENTS

RESULT 1
US-09-977-269-1
Sequence 1, Application US/09977269
Patent No. US20020082037A1
GENERAL INFORMATION:
APPLICANT: DILLRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
APPLICANT: SURES, IRMINGARD
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1260
CURRENT APPLICATION NUMBER: US/09/977,269
PRIORITY FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232,545
PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2000
TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
NAME/KEY: CDS
LOCATION: (258)..(1778)
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte

OTHER INFORMATION: kinase 1
US-09-977-269-1

Alignment Scores:

Pred. No.:	1,05e-266	Length:	2000
Score:	2671.00	Matches:	507
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-977-260-2 (1-507) x US-09-977-269-1 (1-2000)

```

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
DB 258 ATGGCGGGGGGAGGAGCTCTGTTCTCGCGGGGCAATTCACGGCTGTGATTCGTGCGAG 317
QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla 40
DB 318 GAACCTTCCCGGGGTGAGCCCCCTTCTCCGAGCCCTGCGACCCCTCCCTGCTCAGCC 377
QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysTleThrIlyCysGluHisThr 60
DB 378 AGGATGCCAAGAGGGGCTGGGCGCCCGGCGACCCAGTGTATCCAAATCCGAGCACACC 437
QY 61 ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValAlaThrIleLeuGluAla 80
DB 438 CGCCCAAGCCAGGGGAGCTGGGCTTCCGCAAGGGGAGCTGGTCAACCATCTCGAGGCC 497
QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGlyLeuLeu 100
DB 498 TGGGAGAACAGAGCTGTGACCGCGTCAACACCAACAGTGTGACAGAGGGGCTGCTG 557
QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
DB 558 GCAGCTGGGGGCTGGGGGAGGGAGGCCCTCTCCGACAGCCCAAGCTCAGCTCATG 617
QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
DB 618 CCTGTGTTCCACGGGAGATCTGGGCGCAGAGAGCTGTCCAGCAGCTGACGCTCCGAG 677
QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValIleCysVal 160
DB 678 GATGGGCTGTCTGTGGGCGGAGTCCGCGCCGCCACCCCGGCACTACGCTGTGCTG 737
QY 161 SerPheGlyArgAspValIleHisTyrArgValIleHisArgAspGlyHisLeuThrIle 180
DB 738 AGCTTTGGCGCGGAGCTCATCCACTACCGGCTGTGACCGGCGACCCCTCAGCATTC 797
QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys 200
DB 798 GAGGAGCGCGTGTCTTCTGCAACCTCATGACATGTTGAGCATTAACAGCAAGCAAG 857
QY 201 GlyAlaIleCysThrIlySerLeuValArgProLysArgLysHisGlyThrIlySerAlaGlu 220
DB 858 GGGGCTATCTGCACCAAGCTGTGAGCCAAAGCGGAAACCGGAGCAAGTGGCGGAG 917
QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrIleGlnGlyAlaGln 240
DB 918 GAGGAGCTGGCGAGGGGCGGCTGTACTGAACCTGAGCATTTGACATTGGGAGCACAG 977
QY 241 IleGlyGluGlyLeuPheGlyAlaValIleGlnGlyGluTrpLeuGlyGlnLysValAla 260
DB 978 ATGGGAGAGGAGAGTGTGAGCTGTCTGACAGGCTGAGTACTGGGCGCAAAAGTGGCC 1037
QY 261 ValIlyAsnIleLysCysAspValIleThrAlaGlnAlaPheLeuAspGluTrpAlaValMet 280
DB 1038 GTGAGAGATATTCAGAGTGTGAGACAGCCAGGCTTCTTGACAGAGAGCGCGTCATG 1097
QY 281 ThrLysMetGlnHisGlnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
DB 1098 ACGAAGATGCAACAGCAAGAACCTGTGTCTCTGCGTGTATCTCGACACAGGGGCTG 1157

```

```

QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgTrpArgLys 320
DB 1158 TACATGTGATGAGAGCAGTGTGAGCAAGGGCAGAACCTGGTGAACCTTCTGCGGACCCGGGGT 1217
QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
DB 1218 CGAGCCCTCTGTAACCCGCTGACGCTCTGCAAGTCTTCTGTGACGTGGCGGAGGCATG 1277
QY 341 GluTrpLeuGlnSerLysLysLeuValHisArgAspLeuAlaIleArgAsnIleLeuVal 360
DB 1278 GAGTACCTGGAGAGCAAGAAAGCTTGTGACCCGCGGACTGGCCCCCGCAACATCTGTGTC 1337
QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
DB 1338 TCAGAGAGCTGTGGGCAAGGTACAGAGACTTGTGGCTGTGCGCAAGCGGAGCGGAGGGG 1397
QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
DB 1398 CTGAGCTCAAGCCGCGTGGCGGTCAGTGAAGTGACGGCGCCGAGGCTCTCAAAACAGGGAG 1457
QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
DB 1458 TTCACACAGCAAGTCGATGTCTGAGTTTGGGGTGTCTGTGGAGAGTCTTCTCATAT 1517
QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
DB 1518 GCAAGGCGTCCGTAACCTTAAGTCACTAAAGAGGTGTGAGAGCGCGGAGGAAGGGG 1577
QY 441 TyrArgMetGlnProProGlnGlyCysProGlyProValHisValLeuMetSerCys 460
DB 1578 TACCGATGGAACCCCGCGGGGCTGTCCAGGCCCGGTGACGTCTCATGACACCTGTG 1637
QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
DB 1638 TGGGAGGAGAGAGCGCGCGCGCGGCGCACCTTCCGCAACTGGCGAGAAAGTGGCGCG 1697
QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
DB 1698 GAGCTACGCAAGTGTGAGTGGCGCGGCTCGTCTCAGAGGAGAGACCGCGAGCTTCAC 1757
QY 501 SerProArgSerGlnGluPro 507
DB 1758 TCGCCCCGAAAGCCAGAGGCC 1778

```

RESULT 2

US-09-977-260-1
Sequence 1, Application US/09977260
Publication No. US20020192790A1

GENERAL INFORMATION:

APPLICANT: ULLRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
APPLICANT: SURES, IRMINGARD
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1260
CURRENT APPLICATION NUMBER: US/09/977, 260
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232, 545
PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 2000

TYPE: DNA

ORGANISM: Unknown Organism

FEATURE:

NAME/KEY: CDS
LOCATION: (258)..(1778)

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
OTHER INFORMATION: kinase 1
US-09-977-260-1

Alignment Scores:

Pred. No.: 1.05e-266 Length: 2000
 Score: 2671.00 Matches: 507
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-977-260-2 (1-507) x US-09-977-260-1 (1-2000)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 DB 258 ATGGGGGGGAGGCTCTGGTTCTCGGGGGGCAATTTCAGGGCTGATTCGTGAG 317
 QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProValSerAla 40
 DB 318 GAATCTCCCGGGGAGAGCCCGGCTCTCCGAGCCTTGACCCCGCTCCGCTCAGCC 377
 QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGluHisThr 60
 DB 378 AGGATGCCAAGAGGCGCTGGGCGCCGCGGACCCAGTGTATCACCMAATGCGACACACC 437
 QY 61 ArgProLysProGlyGluLeuAlaPheArgGlyGlyAspValValThrIleLeuGluAla 80
 DB 438 CGCCCAAGCCAGGAGGAGCTGGCTCCGCAAGGGGCACTGTGTACCATCTCGAGGCC 497
 QY 81 CysGluAsnLysSerTrpTrpArgValLysHisHisThrSerGlyGlnGlyLeuLeu 100
 DB 498 TGGCAGAACAGAGACTGTTACCGGCTCAGACACACACACATGAGAGAGGGGCTG 557
 QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 DB 558 GCACCTGGGGGCGTGGCGGAGCGGAGGCGCTCTCCGACAGCCCAACCTCAGCCTCATG 617
 QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuInProProGlu 140
 DB 618 CCGGTGTTCCAGGGAGAAATCTCGGGCCAGGAGGCTGCCAGACCTGACCTCCCGCAG 677
 QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuLysVal 160
 DB 678 GATGGCGCTCTCTGGTGGGAGAGTCCCGCGGCCACCCCGGCGACTGCTGTGGCTG 737
 QY 161 SerPheGlyArgAspValIleHisIleTyArgValLeuHisIleAspGlyHisLeuThrIle 180
 DB 738 AGCTTTGGCCGCGAGCTATCCACTACCGGCGTGGTGCAGCGCGAGGCGCACCCTCAGATC 797
 QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisIleTySerLysAspLys 200
 DB 798 GATGAGGCCGTGTCTCTGCAACCTCATGTGACATGGTGAGCATTTACAGCAAGGACAG 857
 QY 201 GluAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 DB 858 GGGCGCTATCTGCACCAAGCTGTGAGACCAAGCGGAACACGGGACCAAGTCGGCCAG 917
 QY 221 GluGluLeuAlaArgAlaGlyTrpPheLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 DB 918 GAGGAGCTGGCCAGGCGGGCTGTTACTGAACCTGCAGCATTTGACATTTGGAGCAAG 977
 QY 241 IleGlyGluGlyGluPheGlyAlaValLeuGlnGlyLysIleTyLeuGlyGlnLysValAla 260
 DB 978 ATCGAGAGGAGAGATTTGGAGCTGTCTGCAAGGTGAGTACCTGGGGGCAAAAGGTGGCC 1037
 QY 261 ValLysAsnIleLysCysAspValThrAlaGlnIlePheLeuAspGluThrAlaValMet 280
 DB 1038 GTGAAGATATCAAGTGTGATGTGACACCCAGGCGCTTCTCGAGGAGACAGCGGCTCATG 1097
 QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyAlaIleLeuHisGlnGlyLeu 300
 DB 1098 ACGAAGATGCAACAGAGACCTGTGTGCTCTCGGCGGTGATCTTCACACAGGGGCTG 1157
 QY 301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 DB 1158 TACATTTGATGAGACAGTGAAGGCAAACTGTGTGATCTTCGCGGAGCCGGGGT 1217

QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGluGlyMet 340
 DB 1218 CGAGCCCTCGTGAACACCGGCTGACACTCTCTGCAAGTCTTCTGCAACGGCGAGGGCATG 1277
 QY 341 GluTrpLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 DB 1278 GAGTACCTGGAGACAGAAAGCTGTGACCGGCACTGGCGCCCGCCCAACATCTCTGTC 1337
 QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
 DB 1338 TCAGAGACCTGTGGCCCAAGGTGACGACTTGTGGCTGGCCAAAGCCGAGGAAAGGG 1397
 QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 DB 1398 CTAGACTCAAGCCGCGTCCCGTCAATGAGAGCGGCCGAGGCTCTCAAAACAGGGAG 1457
 QY 401 PheTrpSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyr 420
 DB 1458 TTCACACAGCAAGTGGATGTTCTGGAGTTTGGGGTGTCTGTGGAGGTTCTTCTCATAT 1517
 QY 421 GlyArgAlaProTrpTrpProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 DB 1518 GGAGGGCTCCGTACCTTAATGTCACTGAAGAAGGTGTGCGAGGCCGTGGAGAGAGGG 1577
 QY 441 TyrArgMetGluProProGluGlyCysProGlyProValHisValLeuMetSerSerCys 460
 DB 1578 TACCGCATGGAACCCCGGAGAGGCTGTGCCAGGCCCGGTGCATCATAGCAGCTGC 1637
 QY 461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
 DB 1638 TGGGAGCGAGAGCCCGCGCGGCCGACCCCTCCGCAAACTGGCGAGAACTGGCCCG 1697
 QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 DB 1698 GAGCTAGCAGTGAAGGTGGCCCGGCTCGTCTCAAGGCGAGAGCGGAGGCTCCACC 1757
 QY 501 SerProArgSerGlnGluPro 507
 DB 1758 TCGCCCGCAAGCCAGAGGCC 1778

RESULT 3

US-09-977-261-1
 ; Sequence 1, Application US/09977261
 ; Publication No. US20030054527A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ULRICH, AXEL
 ; APPLICANT: GISHIKY, MIKHAEL
 ; APPLICANT: SORES, IRMINGARD
 ; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
 ; FILE REFERENCE: 038602/1259
 ; CURRENT FILING DATE: 2001-10-16
 ; PRIOR APPLICATION NUMBER: 08/232,545
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 2000
 ; TYPE: DNA
 ; ORGANISM: Unknown Organism
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (258)..(1778)
 ; FEATURE: (258)..(1778)
 ; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
 ; OTHER INFORMATION: kinase 1
 US-09-977-261-1

Alignment Scores:
 Pred. No.: 1.05e-266 Length: 2000
 Score: 2671.00 Matches: 507
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.008 Indels: 0
 DB: 11 Gaps: 0
 US-09-977-260-2 (1-507) x US-09-977-261-1 (1-2000)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20
 DB 258 ATGGCGGGGGAGGCTCTGCTGGTTCCTGGCGGCAATTCACGGCTGTGATTCTGTGAG 317

QY 21 GluLeuProArgValSerProArgPheLeuArgAlaTrpHisProProProValSerAla 40
 DB 318 GAACCTCCCGGGGTGAGCCCCCTTCCTCCAGCTGGACACCCCTCCCGCTAGCC 377

QY 41 ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrIlyCysGluHisThr 60
 DB 378 AGCATCCCAACGAGCGCTGGGCCCCGGGCACCACTGATACCAAAATGCGAGCACCC 437

QY 61 ArgProLysProGlyLysLeuAlaPheArgGlyGlyAspValValThrIleLeuGluAla 80
 DB 438 CGCCCAAGCGAGGAGCTGGCTTCCTCCAGAGGAGCTGGTCACTCCCTGGAGGCC 497

QY 81 CysGluAsnLysSerTrpTrpArgValLysHisIleThrSerGlyGlnGlyLeuLeu 100
 DB 498 TGGGAACAAGAGCTGATACCCCTCAAGCACACACAGTGGACAGAGGGGCTGCTG 557

QY 101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
 DB 558 GCAGCTGGGGCGCTGGCGGAGCGGAGGCCCTCCGAGACCCCAAGCTCAGCCCTCAG 617

QY 121 ProTrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProProGlu 140
 DB 618 CCTGGTTCACGAGAGATCTGGGCGAGAGGCTGTCCAGAGCTGCGACCTCCCGAG 677

QY 141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTrpValLeuCysVal 160
 DB 678 GATGGGCTGTCTCTGGTGGGAGTCCGGCGCCACCCCGGACACGCTGCTGCTGCTG 737

QY 161 SerPheGlyArgAspAlaIleHisIleTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
 DB 738 AGCTTTGGCGGCGACCTCATCCACTACCGCTGCTCACCAGCGCCACCTCAGCAATC 797

QY 181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisIleSerLysValSPLys 200
 DB 798 GATGAGCGCGTGTCTTCTTCGCAACCTCATGAGCATGTGGACATTACAGAGCAAG 857

QY 201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerAlaGlu 220
 DB 858 GGGGCTATCTGCACCAAGCTGGTGAGACCAAGCGGAACAGCGGAGCAAGTGGCGAG 917

QY 221 GluGluLeuAlaArgAlaGlyTrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGln 240
 DB 918 GAGGACTGGCCGAGGGCGGCTGCTGTAACCTCAGCATTTGACATTGGGAGACAG 977

QY 241 IleGlyGluGlyValPheGlyAlaValLeuGlnGlyValLeuThrLeuGlyGlnLysValAla 260
 DB 978 ATCGGAGAGGAGAGTGTGGAGCTGTCTGCGAGGAGAGCTGGGGCAAAAGGTGGCC 1037

QY 261 ValLysAsnIleLysCysAspValThrAlaGlnAlaPheLeuAspGluThrAlaValMet 280
 DB 1038 GTCAAGAATATCAAGTGTATGTGACAGCCAGGCTTCTCTGACAGAGCGCGGTGAG 1097

QY 281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
 DB 1098 ACAGAAATGCAACACAGAAACCTGCTGCTCTGCTGGGCTGATCTCGACACAGGGGCTG 1157

QY 301 TyrIleValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
 DB 1158 TACATTGTATGAGGACGTGAGCAAGGGCAACCTGTGAACTTTCTGGGAGACCCGGGCT 1217

QY 321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMet 340
 DB 1218 CGAGCGCTGTGACACCGCTCAGCTCTGCACTTTCTCTGACAGTGGCCGAGGAGCATG 1277

QY 341 GluTrpLeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
 DB 1278 GAGTACCTGGAGAGCAAGCTTGTGCAACCGGACCTGGCCCGCCGCAACATCTGTGCTC 1337

QY 361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380
 DB 1338 TCAGAGAGACCTGGTGGCCAGAGGTGACAGACTTGGCTGGCCAAAGCCGAGCGAAGGGG 1397

QY 381 LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
 DB 1398 CTAGACTCAAGCGGCTGGCCGCTCAAGTGGAGCGGCCCGAGGCTCTCAAAACAGGGGAG 1457

QY 401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrp 420
 DB 1458 TTCACCAAGCAAGCGGATGTGTGGAGTCTGGGCTGCTCTGGGAGGCTTCTCATAT 1517

QY 421 GlyArgAlaProTrpProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
 DB 1518 GAGCGGCTCCGTAACCTTAATGTCATGAAGAGGTGTGGAGGCCGTGGAGAGGGG 1577

QY 441 TyrArgMetGluProProGluLysCysProGlyProValHisValLeuMetSerSerCys 460
 DB 1578 TACCGATGGAACCCCGCGAGGCTGTCCAGGCGCCGCTGACAGTCTCATAGACAGCTGC 1637

QY 461 TrpGluAlaGluProAlaArgArgProProPheArgLysLeuAlaGluLysLeuAlaArg 480
 DB 1638 TGGGAGGAGAGCGCCCGCGCGCCACCTTCCGAAACTGGCCGAGAAAGCTGGCGCGG 1697

QY 481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
 DB 1698 GAGCTAGCAAGTGCAGAGTGCAGGCTGCTGCTGCTCAGGGCAGGAGCCGAGGCTCCACC 1757

QY 501 SerProArgSerGlnGluPro 507
 DB 1758 TCGCCCGGAAGCCAGAGAGCCC 1778

RESULT 4
 US-10-084-817-341
 ; Sequence 341, Application US/10084817
 ; Publication No. US20030119009A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Susan Stuart
 ; APPLICANT: Jedd G. Nuchtern
 ; APPLICANT: Sharon E. Plon
 ; APPLICANT: Jason M. Shohet
 ; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
 ; FILE REFERENCE: PA-0046 US
 ; CURRENT APPLICATION NUMBER: US/10/084, 817
 ; PRIOR FILING DATE: 2002-02-25
 ; PRIOR APPLICATION NUMBER: 60/270,784
 ; NUMBER OF SEQ ID NOS: 365
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 341
 ; LENGTH: 1989
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Inocyte ID No. US20030119009A1 1794861CB1
 US-10-084-817-341

Alignment Scores:
 Pred. No.: 4,07e-261 Length: 1989
 Score: 2617.00 Matches: 505
 Percent Similarity: 99.618 Conservative: 0
 Best Local Similarity: 99.618 Mismatches: 2
 Query Match: 97.988 Indels: 2
 DB: 14 Gaps: 0

US-09-977-260-2 (1-507) x US-10-084-817-341 (1-1989)

QY 1 MetAlaGlyArgGlySerLeuValSerTrpArgAlaPheHisGlyCysAspSerAlaGlu 20

```

Db      259 ATGGGGGGGCGAGGCTCTGTGTTCTCGCGGGCATTTACGGCTGTGATTCTCTGAG 318
QY      21  GIUleuProArgValSerProArpHeuArpAlaTrpHisProProAlaSerAla 40
Db      319 GAACTTCCCGGGGTAGCCCGCGCTCTCCGAGCGTGGACACCCCTCCCTCCGACCC 378
QY      41  ArgMetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCysGlnHisThr 60
Db      379 AGGATGCCACAGAGCGGTGGGGCCCGGGGCAACCGATGTATACCAATTCGACACACC 438
QY      61  ArgProLysProGlyGluLeuAlaPheArgLysGlyAspValThrIleLeuGlnAla 80
Db      439 CGCCCAAGCCAGGAGGAGCTGGCTCCGCAAGGCGGACGTGTACCATCTCGAGAGGCC 498
QY      81  CysGluAsnLysSerTrpTrpArgValLysHisIleThrSerGlyGlnGlnGluLeu 100
Db      499 TCCGGAACACAGAGTGTACCGCGTCAAGCACACACAGTGGACAGAGGGGCTGTG 558
QY      101 AlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMet 120
Db      559 GCAGCTGGGGGCGCTCGGGGAGGGGAGGCCCTCTCCGACAGACCCCAAGCTCAGCTCATG 618
QY      121 ProTrpPheHisGlyLysIleSerGlyGlnGlnAlaValGlnGlnLeuGlnProGlu 140
Db      619 CGGTGTTCACAGGGAAGATCTCGGGCCAGGAGCGTGTCCAGACGTCGACGCTCCGCGAG 678
QY      141 AspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal 160
Db      679 GATGGGCTGTTCCTGTGCGGAGTCCGCGGACCCGCGGAGTACGCTGTGCGTGTG 738
QY      161 SerPheGlyArgAspValIleHisIleTrpArgValLeuHisArgAspGlyHisLeuThrIle 180
Db      739 AGCTTTGGCCCGGAGAGTATCCACACCGCGTGTGCGACCGGAGCGGCACCTCACAATC 798
QY      181 AspGluAlaValPhePheCysAsnLeuMetAspMetValGluHisTyrSerLysAspLys 200
Db      799 GATGAGGCCGCTGTCTGTGCAACCTCATGACATGTGAGCATTTACAGCAAGAGACAAG 858
QY      201 GlyAlaIleCysThrLysLeuValArgProLysArgLysHisGlyThrLysSerLysGlu 220
Db      859 GCGCTATCTGACCAACGCTGGTGGACCAAGCGGAAACAGGGAGCAAGCGGCGGCGAG 918
QY      221 GluGluLeuAlaArgAlaGlyTrpLeuAsnLeuGlnHisLeuThrLeuGlnValGln 240
Db      919 GAGGAGCTGGCCAGGCGGGCGGTGTACTGAACCTGCAGCACTTACACTTGGGACACAG 978
QY      241 IleGlyGlnGlyGluPheGlyAlaValLeuGlnGlyLysTrpLysGlnLysValAla 260
Db      979 ATCGAGAGGAGGAGACTTTGGAGCTCTCTCGCAGGGGTGAGTACTGGGCAAAAGGTGCC 1038
QY      261 ValLysAsnIleLysCysAspValThrAlaGlnIlePheLeuAspGlyThrAlaValMet 280
Db      1039 GTGAAGAAATATCAATGTGATGTGACAGCCCGGCTTCTCTGGAGAAAGCGGCGCTCATG 1098
QY      281 ThrLysMetGlnHisGluAsnLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeu 300
Db      1099 ACGAAGATGCACACAGAACCTGTGTGCTCCGCGGCGTGTCTCTGACACAGGCGCTG 1158
QY      301 TyrIleValMetGluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGly 320
Db      1159 TACATTTCTCATGAGACAGCTGAGCAAGGCAACCTGTGTAACCTTCTCGGACCCGGGGT 1218
QY      321 ArgAlaLeuValAsnThrAlaGlnLeuLeuGlnPheSerLysValAlaGlnGlyMet 340
Db      1219 CGAGCCCTGCTGAACACCGCTCAGCTCTGCGAGTTTCTCTGCACTGTGCGAGGGCATG 1278
QY      341 GluTrpLeuGluSerLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuVal 360
Db      1279 GAGTACCTGTGAGAGCAAGAGCTGTGACCGCGACCTGGCGCGCGCAACTCTGTGTC 1338
QY      361 SerGluAspLeuValAlaLysValSerAspPheGlyLeuAlaLysAlaGluArgLysGly 380

```

```

Db      1339 TCAGAGACCTGCTGCGCCAGAGTTCAGCGACTTTGGCTGCGCCAAAGCCGAGCGAAGGGC 1398
QY      381  LeuAspSerSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLys 400
Db      1399 CTAGACTCAAGCGGCTGCGCTCAAGTGAAGGAGGCGCGAGGCTCTCAAAACACGGGAAG 1458
QY      401 PheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerThr 420
Db      1459 TTCACACAGCAAGTGGAGTGTGAGATTGTGGGGTGTGCTGTGTGGAGAGTCTTCATAT 1518
QY      421 GlyArgAlaProTyrProLysMetSerLeuLysGluValSerGluAlaValGluLysGly 440
Db      1519 GGACGGGCTCGCTACCTTAATAATGTCACTGAAGAAGAGTGTGCGAGGCGCTGGAGAAGGGG 1578
QY      441 TyrArgMetGluProProGlnGlyGlyProGlyProValHisValLeuMetSerSerLys 460
Db      1579 TACCCATGGAACCCCGAGGCGGTGTCCAGGCGCCGCTGACAGTCTCATGTGAGCAGTGC 1638
QY      461 TrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLysLeuAlaArg 480
Db      1639 TGGAGGCGAGAGCGGCC-CGCGGCGCACCTTCCGCAAACTGGCCGAGAGCTGGCCGG 1697
QY      481 GluLeuArgSerAlaGlyAlaProAlaSerValSerGlyGlnAspAlaAspGlySerThr 500
Db      1698 GAGCTACCGCAGTGCAGGTGCGCCAGCGCTCCGTCAGGGGCAAGAGCGCGACGCTCC-ACC 1756
QY      501 SerProArgSerGlnGluPro 507
Db      1757 TCGCCCGAAGCCAGGAGGCC 1777

RESULT 5
US-09-954-531-188
; Sequence 188, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0.
; SEQ ID NO 188
; LENGTH: 2187
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-188

Alignment Scores:
Pred. No.: 3,276-119 Length: 2187
Score: 1247.50 Matches: 238
Percent Similarity: 71.59% Conservative: 82
Best Local Similarity: 53.24% Mismatches: 118
Query Match: 46,718 Indels: 9
Dh: 10 Gaps: 3

US-09-977-260-2 (1-507) x US-09-954-531-188 (1-2187)
QY      40  AlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
Db      116  GCCAGAGTCTCTGAGAGATGTCAAGCAATACAGCGCGCTGTGCGCATTCGGTACAGAAATGT 175

```

QY 54 ILeThrIyScySgLnHstHrArGProLySProGlyGluLeuAlaPheArGlySGLyASP 73
 Db 176 ATGGCCAACTGACAACTTCCAGCGCACTCCGAGAGAGACCTGCCCTTCTCCCAAGGAGAC 235
 QY 74 ValValThrIleLeuGlnAlaCysGluAsnLysSerTrpTyrArgValLysHstHstHr 93
 Db 236 GTGCTCAACATTGTGGCCGTCACCAAGACCACCACTGTACCAAGCCAAAGCAAGAGTG 295
 QY 94 SerGlyGlnGluLeuLeuAlaAlaGlyAlaLeuArgLysGluAlaLeuSerAla 113
 Db 296 ---GGCCGTGAGGACATCATCCAGCACTACCTCCAGAGAGCGGAGCGGTGAAGCGC 352
 QY 114 AspProLySLeuSerLeuMetProTrpPheHisGlyLysLysSerGlyGlnGluAlaVal 133
 Db 353 GGTAACCAACTGACCTCATGCTTGGTTCACGGCAAGATCAACAGGAGCGAGCGTAC 412
 QY 134 GlnGlnLeuGlnProGlnAlaAspGlyLeuPheLeuValArgLysSerAlaArgHstPro 153
 Db 413 CGGCTTGTGTACCGCGCGGAGACAGCGCTTCTGTGGCGGAGAGACCACTACCC 472
 QY 154 GlyAspTyrValLeuCysValSerPheGlyArgAspValIleHstTyrArgValLeuHis 173
 Db 473 GGAGACTACACAGCTGTGCTGAGCTCGAGCGCAAGGTGGAGAGCATCCGCAATCATGTAC 532
 QY 174 ArgAspGlyHstLeuThrIleAspGlnAlaValPhePheCysAsnLeuMetAspMetVal 193
 Db 533 CATGCCAGCAAGCTCAGCATCGACAGAGAGGTACTTGAAGCACTCATAGCACTG 592
 QY 194 GluHstTyrSerLysAspLysGlyAlaIleCysThrLysLeuValArgProLySArgLys 213
 Db 593 GAGCACTACACATCAAGAGCGAGATGGACTGTGTACGGCGCTCATTAACCAAGGTATG 652
 QY 214 HisGlyThrLysSerAlaGluGluLeuAlaArgAlaGlyTyrPheLeuAsnLeuGln 233
 Db 653 GAGGGCACATGGCGGGCGAGATGATTTACCGCGAGCGGTGGCGCTGAACATGAG 712
 QY 234 HisLeuThrLeuGlyAlaGlnIleGlyLysGlnGlnPheGlyAlaValIleGlnGlyGlu 253
 Db 713 GAGCTGAGCTGCTGCAGACCATCGAGAGGGAGAGTTCGAGACGGATGCTGGGAT 772
 QY 254 TyrLeuGlyGlnLysValAlaValLysAsnIleLysCysAspValThrAlaGlnAlaPhe 273
 Db 773 TACCGAGGAAACAAAGTCGCGCTCAAGTCAATTAAGACAGACCCCACTGCCAGCGCTTC 832
 QY 274 LeuAspGlyThrAlaValMetThrLysMetGlnHisGluAsnLeuValArgLeuGly 293
 Db 833 CTGGCTGAAACCTCAGTCATGACAGCAACTCGGCATAGCAACTGGGTGACGCTCGGCG 892
 QY 294 ValIleLeuHstGln-----GlyLeuTyrIleValMetGlnHstLysSerGlyAsn 311
 Db 893 GTGATCGTGAGAGAGAGGGCGGCTCTACATCGTCACTGATGATGAGGCAAGGGAGCG 952
 QY 312 LeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGlnLeuGln 331
 Db 953 CTGTGTGACTACTCGCGTCTAGGGGTGCGTCAAGTGTGGCGGAGACTCTCTCTCAAG 1012
 QY 332 PheSerLeuHstValAlaGluGlyMetGlyTyrLeuGlnSerLysLeuValHisArg 351
 Db 1013 TTCTCCCTAGATGTCTGCGAGGCGCATGATACCTGGAGGCAACAATTTCGTGACATCGA 1072
 QY 352 AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSerAspPhe 371
 Db 1073 GACCTGCTGCGCCGCAATGTGTGTCTGAGAGCAACGTGGCCAAAGGTGACGACTTT 1132
 QY 372 GlyLeuAlaAlaArgLysGlyLysAspSerSerArgLeuProValLysTrpThr 391
 Db 1133 GGTCTACCAAGAGAGCGCTCCAGACCAAGAGACAGCGCAAGCTGCCAGTCAAGTGAC 1192
 QY 392 AlaProGlnAlaLeuLysHstGlyLysPheThrSerLysSerAspValTrpSerPheGly 411
 Db 1193 GCCCTGAGGCGCTGAGAGAGAGAAATTCCTCCACTAAGTCTGACGCTGGAATTCGGA 1252
 QY 412 ValLeuLeuThrPheValPheSerTyrGlyArgAlaProTyrProLySMetSerLeuLys 431

Db 1253 ATCCTTCTGCGGAAACTACTCTTTGGCGAGTGCCTTATCCAAATTCCTCGAAG 1312
 QY 432 GluValSerLysAlaValGluLysGlyTyrArgMetGluProProGlnGlyCysProGly 451
 Db 1313 GACTCTCTCCCTCGGGGTGGAGAGAGGTCTACAAATGATGATGCCCGGAGCGCTGCC 1372
 QY 452 ProValHstValLeuMetSerSerCysTrpGlnAlaGluProAlaArgArgProPhe 471
 Db 1373 GCAGTCTATAGACATGAGAACTGCTGACCTGAGCGCCGCAATGAGCGCTCTTC 1432
 QY 472 ArgLysLeuAlaGluLysLeu 478
 Db 1433 CTACAGCTCCGAGAGCACTT 1453
 RESULT 6
 US-10-298-377A-1
 : Sequence 1, Application US/10298377A
 : Publication No. US20030130209A1
 : GENERAL INFORMATION:
 : APPLICANT: The Scripps Research Institute
 : APPLICANT: Chersesh, David A.
 : APPLICANT: Paul, Robert
 : APPLICANT: Elliceiri, Brian
 : TITLE OF INVENTION: Method of Treatment of Myocardial
 : TITLE OF INVENTION: Infarction
 : FILE REFERENCE: TSRI-651.5
 : CURRENT APPLICATION NUMBER: US/10/298, 377A
 : CURRENT FILING DATE: 2002-11-18
 : PRIOR APPLICATION NUMBER: 10/298, 377
 : PRIOR FILING DATE: 2002-11-18
 : PRIOR APPLICATION NUMBER: 09/470, 881
 : PRIOR FILING DATE: 1999-12-22
 : PRIOR APPLICATION NUMBER: 09/538, 248
 : PRIOR FILING DATE: 2000-03-29
 : PRIOR APPLICATION NUMBER: PCT/US99/11780
 : PRIOR FILING DATE: 1999-05-28
 : PRIOR APPLICATION NUMBER: 60/087, 220
 : PRIOR FILING DATE: 1998-05-29
 : NUMBER OF SEQ ID NOS: 4
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1
 : LENGTH: 2187
 : TYPE: DNA
 : ORGANISM: homo sapiens
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (134)...(1486)
 : US-10-298-377A-1
 Alignment Scores:
 : Pred. No.: 3,27e-119 Length: 2187
 : Score: 1247.50 Matches: 238
 : Percent Similarity: 71.59% Conservative: 82
 : Best Local Similarity: 53.24% Mismatches: 118
 : Query Match: 46.71% Indels: 9
 : DB: 15 Gaps: 3
 US-09-977-260-2 (1-507) x US-10-298-377A-1 (1-2187)
 QY 40 AlaArgMetProThrArg-----ArgTrpAlaProGlyThrGlnCys 53
 Db 116 GCCAGAGCTCTCAGAAAGATGACAGAAATACAGCGCGCTTGTGCAAGAGATGT 175
 QY 54 ILeThrIyScySgLnHstHrArGProLySProGlyGluLeuAlaPheArGlySGLyASP 73
 Db 176 ATGGCCAACTGACAACTTCCAGCGCACTCCGAGAGAGACCTGCCCTTCTGCAAGAGAGAC 235
 QY 74 ValValThrIleLeuGlnAlaCysGluAsnLysSerTrpTyrArgValLysHstHstHr 93
 Db 236 GTGCTCAACATTGTGGCCGTCACCAAGACCACCACTGTACCAAGCCAAAGCAAGAGTG 295
 QY 94 SerGlyGlnGluLeuLeuAlaAlaGlyAlaLeuArgLysGluAlaLeuSerAla 113

```
Db 296 ---GGCGCTGAGGCGATCCACGACAGTCTGCGAAGACGGGAGCGCTGAGGG 352
Qy 114 ASPRPLYSLSESLMEETPTTPRPHENISGLYSLISESGVLGGLVAL 133
Db 353 GGTACCAACATCGAGCTCATGCTTGTTCCAGCGCAAGATCATACGGGAGCGCTGAG 412
Qy 134 GINGLLEUGLNPProGluAPSLLeuPheLeuValArgGluSerAlaArgHisPro 153
Db 413 CGGCTTCTGTACCCCGGAGACAGGCTGTTCTGTGTCGGGAAGACCCACTACCC 472
Qy 154 GLYASPTyrValLeuCySValSerPheGlyArgAspValIleHisTyrArgValLeuHis 173
Db 473 GGAGACATACAGCTGTGGTGAGTCGCGACGGCAAGGTGAGACATACCGCATCATGTCAC 532
Qy 174 ARGASRGLYHISLeuThrIleAspGluAlaIlePhePheCySAsnLeuMetLaspMetVal 193
Db 533 CATGCCACCACTAGATCGACAGGAGGAGGTGACTTTGAGAACCTCATGCGAGCTGTG 592
Qy 194 GIUHSYrSerLysAspLysGlyAlaIleCySThrLysLeuValArgProLysArgLys 213
Db 593 GAGCACTACCTCAGACGACAGATGACTCTGTACGGCTCATTAACCAAGGTCATG 652
Qy 214 HISGLYThrLysSerAlaGluGluIleuValArgAlaGlyTyrLeuLeuAsnLeuGln 233
Db 653 GAGGCGACAGTGGCGCCAGAGATGATTCACCGCAGCGCTGGCGCCCTGAACATGAAAG 712
Qy 234 HISLeuThrLeuGlyAlaGluIleGlyGluGlyPheGlyAlaValLeuGlnGlyGlu 253
Db 713 GAGCTGAAGCTCTCTCAGACACATCGGGAAGGGGAGTTCGGAGCTGTGCTGGCGCAT 772
Qy 254 TYrLeuGlyGluIleValAlaValAlaValLysAsnIleLysCySAspValThrAlaGlnLaphe 273
Db 773 TACCGAGGAGACAAAGTGGCCCTCAAGTGCATTAAGAACGACGCCACTGCCAGGCTTTC 832
Qy 274 LeuAspGlyThrAlaValMetThrLysMetGlnHisGluAsnLeuValArgLeuLeuGly 293
Db 833 CTGGCTGAAGCTCATGATGACGACACTGCGCATAGACACCTGGTGCAGCTCTGGGCG 892
Qy 294 VALIleLeuHisGln-----GlyLeuTyrIleValMetGlnHisValSerLysGlyAsn 311
Db 893 GTGATCTGTGGAAGAGCGCGGCTCATCATCGTACATCGTACATGATCATGCGCAAGGGAGAC 952
Qy 312 LeuValAsnPheLeuArgThrArgGlyArgAlaValAlaValAspThrAlaGlnLeuGln 331
Db 953 CTGTGGAGCTACCTCGGCTGAGGGTGGTGCAGTCTGGGGGAGACTGTCTCTCAAG 1012
Qy 332 PheSerLeuHisValAlaGluGlyMetGlyTyrLeuGlnSerLysLysLeuValHisArg 351
Db 1013 TTCTGCTAGATGTCTGCGAGCGCATGGAATGACTGAGAGGCAACAATTTCTGCATCGA 1072
Qy 352 ASPLeuAlaIleArgAsnIleLeuValSerGluAspLeuValAlaLysValSerPhe 371
Db 1073 GACCTGTGGTCCCGGAAATGCTGTGGTGTGAGACAACTGAGGCCAAGGTCAGCACTTT 1132
Qy 372 GYLLeuAlaLysAlaGluArgLysGlyLeuAspSerSerArgLeuProValLysTrpThr 391
Db 1133 GGTCTCACCAAGAGAGCGGTCCAGCAACCCAGACAGGGGCAACTGCGCATCAAGTGGACA 1192
Qy 392 ALAPROGluAlaLeuHisGlyHisGlySerPheThrSerLysSerAspValTrpSerPheGly 411
Db 1193 GCCCTGAGGCGCTGAGAGAAATGCTCCACTAAGTGTGAGCGTGTGAGATTTCGGA 1252
Qy 412 VALIleLeuThrGluValAlaPheSerTyrGlyArgAlaProTyrProLysMetSerLeuLys 431
Db 1253 ATCTCTCTGTGGAAATCTACTCTTTGGCGAGTCCCTTATCCAAATTTCCCTCCGAG 1312
Qy 432 GIUValSerGluAlaValAlaGlyLysGlyTyrArgMetGluProProGluGlyLysProGly 451
Db 1313 GACGTGTGCTCGGTGTGAGAGAGGCTCAAGATGAGATGCCCGCCAGCGCTGCGCGCC 1372
Qy 452 ProValHisValLeuMetSerSerCyStrpGluAlaGluProAlaArgArgProPhe 471
```

```
Db 1373 GCAGTCTGATGATCATGAGAACTGTGGACACTGAGCGCGCCATCGGCGCTTC 1432
Qy 472 ARGLYSLLeuAlaGluLysLeu 478
Db 1433 CTACAGTCTCGAGAGCAGCTT 1453

RESULT 7
US-10-177-293-87
; Sequence 87, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatz, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, VIC
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzsati, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 2420
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-87

Alignment Scores:
Pred. No.: 3,76e-119 Length: 2420
Score: 1247.50 Matches: 238
Percent Similarity: 71.598 Conservative: 82
Best Local Similarity: 53.248 Mismatches: 118
Query Match: 46.71% Indels: 9
DB: 14 Gaps: 3

US-09-977-260-2 (1-507) x US-10-177-293-87 (1-2420)

Qy 40 ALAArgMetProThrArg-----ArgTyrAlaProGlyThrGlnCyS 53
Db 395 GCCAGAGCTCTGAGAAAGATGTGCAGCAATACAGCGCGCTGCGCAATCCGTCAGAGATGT 454
Qy 54 IIEThrLysCySGLNHisThrArgProLysProGlyGluLeuAlaPheArgLysGlyAsp 73
Db 455 ATTCGCAAGTACAACTTCCAGCGCACTGCGGAGCAAGAGACAGTCCCTCTGCAAGAGAC 514
Qy 74 VALValThrIleLeuGluAlaCySGLNHisThrArgValArgValHisThr 93
```

```

Db 515 GTGCTCAACATTTGGCGGTGACACAGCCCAAGTGTCAACAAACCAAGAGTGC 574
Qy 94 SerIleuGlnGluGlyLeuLeuAlaAlaGlyAlaLeuArgGluArgGluAlaLeuSerAla 113
Db 575 ---GGCGGTGAGGCGATCATCCAGCAACATGCTCCAGAGCGGAGGCGGTGAAGCGC 631
Qy 114 AspProIysLeuSerLeuMetProTyrPheHisGlyIstIleSerGlyGlnGluAlaVal 133
Db 632 GGTACCAAACTCAGCTCATGCTGTGGTTCACGGCAAGATCAACAGCGGACAGCTGAG 691
Qy 134 GlnGlnLeuGlnProGluAspGlyLeuPheLeuValArgGluSerAlaArgHisPro 153
Db 692 CGGCTCTGTACCCGCGGAGACAGGCTGTCTGTGGCGGAGACACCACTACCC 751
Qy 154 GlyAspTyrValLeuCysValSerPheGlyArgAspValIleHisTyrArgValLeuHis 173
Db 752 GGAACATCAACAGCTGTGCTGAGCTGCGAGCGCAAGGTGGAGCACTACCGCATCATGATC 811
Qy 174 ArgAspGlyHisLeuThrIleAspGlyAlaValPhePheCysAsnLeuMetAspMetVal 193
Db 812 CATGCCAGCAAGCTCAGCATGACGAGAGGTGTACTTTGAGAACCTCATGCAAGTGTG 871
Qy 194 GlnHisTyrSerIysAspIysGlyAlaIleCysThrIstIstLeuValArgProIysArgLys 213
Db 872 GAGCATACACACTCAGACGAGAGATGAGCTGTACGGCGCTCATTAACCAAGGTGATG 931
Qy 214 HisGlyThrIstSerAlaGluGluGlnLeuAlaArgAlaGlyTyrPheLeuAsnLeuGln 233
Db 932 GAGGGCACAGTGGCGGCGGAGATGATTTACCGCGAGCGGCGGCGCTGCAACATGAG 991
Qy 234 HisLeuThrLeuGlyAlaGlnIleGlyGluGlyGlnPheGlyAlaValIleGlnGlyGlu 253
Db 992 GAGCTGAAGTGTCTGACGACCATCGGAGAGGAGGTTCGGAACAGTGAAGTGTGGCGAT 1051
Qy 254 TyrLeuGlyGlnIstValAlaValIstAsnIleCysCysAspValThrIleGlnAlaPhe 273
Db 1052 TACCGAGGGAACAAAGTCCGCGTCAAGTGAAGACACACCCACTGCCAGGCTTC 1111
Qy 274 LeuAspGlyThrAlaValMetThrIstMetGlnHisGlnAsnLeuValArgLeuGly 293
Db 1112 CTGGCTGAACCTCAGTACGAGCAACTGCGCATGCAACTGGTGCGAGTCTCGGCGC 1171
Qy 294 ValIleLeuHisGln-----GlyLeuTyrIleValMetGlnHisValSerIstGlyAsn 311
Db 1172 GTGATGTGGAGGAGAAAGGCGGCTGTACATGCTGACTGAGTACATGCGCAAGGAGGAGC 1231
Qy 312 LeuValAsnPheLeuArgTyrArgGlyArgGlyAlaLeuValAsnThrIleGlnLeuGln 331
Db 1232 CTGTGTGACTACCTGGGTCTAGGGTGGTGTGCTGTGCGGAGACATGCTCTCTCAAG 1291
Qy 332 PheSerLeuHisValAlaGluGlyMetGluTyrLeuGluSerIstIstValHisArg 351
Db 1292 TTCCTGCTGATGATTCGCGAGGCGCATGATACCTGAGGAGCAACAATTCGTGATCGA 1351
Qy 352 AspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaIstIstValSerAspPhe 371
Db 1352 GACCTGCGTCCCGCATGCTGTGTGTGAGGAGCAAGTGGCGCAAGGTCAGCACTTT 1411
Qy 372 GlyLeuAlaIstAlaGluArgIstGlyLeuAspSerSerArgLeuProValIstPThr 391
Db 1412 GGTCTCACCAGAGGCGTTCACGACCAAGACAGGAGAGTGCACATGCAAGTGGACA 1471
Qy 392 AlaProGluAlaLeuIstIstGlyIstPheThrSerIstSerAspValIstPheGly 411
Db 1472 GCCCGTAGGCGCTGAGAGAGAAATTCACATAAGTCACTGTGAGATTTCGGA 1531
Qy 412 ValLeuLeuTyrGlnValPheSerTyrArgAlaProTyrProIstMetSerLeuLys 431
Db 1532 ATCCTTCTGTGGAAATCTACTTCTGTGGCGAGTGCATTATCCAAAGTTCCTCCGGAAG 1591
Qy 432 GluValSerGluAlaValGluIstGlyTyrArgMetGluProProGluGlyCysProGly 451

```

```

Db 1592 GACGTGCTCCCTCGGCTGGAGAGGCTACAAAGATGATGCCCGGAGCGCTCCCGCC 1651
Qy 452 ProValHisValLeuMetSerCysTyrPheGluAlaGluProAlaArgArgProPhe 471
Db 1652 GCACGTATAGATGATGATGAGAACTGCTGGCACTGAGAGCGGCGCATGCGGCTCTTC 1711
Qy 472 ArgIstLeuAlaGluIstIstLeu 498
Db 1712 CTACAGCTCCGAGAGCACTT 1732

RESULT 8
US-09-954-456-1983
: Sequence 1983, Application US/09954456
: Patent No. US20020115057A1
: GENERAL INFORMATION:
: APPLICANT: Young, Paul
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954, 456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233, 617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234, 052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234, 923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235, 134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235, 637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235, 638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235, 711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235, 720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235, 840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235, 863
: PRIOR FILING DATE: 2000-09-27
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: Patent version 3.0
: SEQ ID NO 1983
: LENGTH: 2015
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-954-456-1983

Alignment Scores:
Pred. No.: 3,4e-65 Length: 2015
Score: 725.50 Matches: 180
Percent Similarity: 51.25% Conservative: 87
Best Local Similarity: 34.55% Mismatches: 187
Query Match: 27.16% Gaps: 67
DB: 10 Gaps: 15

US-09-977-260-2 (1-507) x US-09-954-456-1983 (1-2015)
Qy 34 HisProProValSerIstAlaArg----- 41
Db 272 CATCCACCA-----TCAAGCCGCGGCTTAATAGCCACAGCAACACACAGCAATCA 325
Qy 42 -----MetProThrArgTyrPheAlaProGlyTyrGlnCysIleThrIstIstCys 57
Db 326 GGGAGGACAGCGCTGAGAGCAATCAATGCTGTGCCC-----TGT 364
Qy 58 GlnHisThrArgPro-LysPro-GlyGlnLeuAlaPheArgIstGlyAspValIstPThrI 77
Db 365 ATGATTACAGGCGCATTCACACAGAAAGCTGAGTTCAGAAAGGAGGAGCAGATGCTGG 424
Qy 77 IsteGluAlaIstCysGluAsnIstSerTyrArgValIstIstIstIstSerGlyGln 97

```

```

Db 425 TCCTAGAG-----GAATCCGGGAGTGTGGGAAGGCTCGATCCCTGGCCACCCGGGAAG 478
Qy 1uglyLeuLeuAlaIaglyAlaLeuArgIuArgIuAlaLeuSerAlaAspProlySL 117
Db 479 AGGGCTACATCCCAAGCAACTATGTCCGCCGTTGACTCTCTGGAGACAG-----531
Qy 117 euserLeuMetProTPrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuG 137
Db 532 -----GAATGGTTTTCAGAGGCAATCAGCCGGAAGAGGAGGAGCCGCACTGC 580
Qy 137 InProProGluAsp-----GlyLeuPheLeuValArgGluSerAlaArgHisProGlyA 155
Db 581 TGCGTCCCGGCAATGCTGGGCTCTCATGATCCGGATACCGAGACACCTAAAGGAA 640
Qy 155 spTyValLeuCyVal-----SerPheGlyArgAspValIleHisTyra 170
Db 641 GCTACTCTTGTCCGTGGAGACTAGACCTCCGCAAGGAGATACCGTGAACATTCACA 700
Qy 170 rGValLeuHisArgAsp---GlyHisLeuThrIleAspGluAlaValPhePheCysAsnL 189
Db 701 AGATCCGGAGACCTCGACACAGGGGGCTTTCATATCCCGCCGAGACCTTCAGACCTC 760
Qy 189 euserMetValGlnHisTySerLysAspLysGlyAlaIleCysThrLysLeuValA 209
Db 761 TGCAAGAGCTGTGTGACCACTACAAGAGGGAAGGAGGAGGAGGAGGAGGAGGAGGAG 820
Qy 209 rGrProLysArgGlyHisGlyLysSerAlaGluGlnGluAlaArgValArgLysTrpL 229
Db 821 TGCCCTGCATG-----TCTTCCAGCCCGCAAGAGCTTGGAGAGAAATGCTCTGG 871
Qy 229 euserAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyLysGlyLeuValArg 249
Db 872 AGATCCCTCGGGAGATCCCTCAAGCTGGAGAGAAACTTGGAGCGGGAGTGTGGGAG 931
Qy 249 alleuGlnGlyLysTy---LeuGlyLysValAlaValLysAsnLys---CysA 267
Db 932 TCTGATAGGCCCACTCAACAGCAAGCAACCAAGGTGGGAGTGAAGTGAAGCAGAGGA 991
Qy 267 spValThrAlaGlnAlaPheLeuAspGluThrAlaValMetThrLysMetGlnHisGlu 287
Db 992 GCATGTGCGTGGAGGCTTCTCGCAGAGGCCAAGGATGATAAACTGTGAGCATGCA 1051
Qy 287 snLeuValArgLeuGlnGlyValIleLeuHisGlnGlyLysTrpLysLeuValMetGlnHis 307
Db 1052 AGCTGGTCAAACTTCATGCGGTGGTCAACAGAGGCCATTCATCATCATCAGGAGTTCA 1111
Qy 307 alSerLysGlyAsnLeuValAsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrA 327
Db 1112 TGCCCAAGGAAGGCTGTGCTGACTTCTGAAAGTGAAGGAGGAGCAGCAGCATTCG 1171
Qy 327 lagInLeuLeuGlnPheSerLeuHisValAlaGlnGlyMetGlnTyrlLeuGlnLysTrpL 347
Db 1172 CAAACTCATGACTTCTCAGCCAGATTCAGAAAGGATGCGCTTCATTCAGAGAGGA 1231
Qy 347 yslLeuValHisArgAspLeuAlaIaArgAsnIleLeuValSerGlnLysAspValAla 367
Db 1232 ACTACATCCACCGAGACTCCGAGCTGCAACATCTTGCTGTGATCCCTGGTGTCTA 1291
Qy 367 yslValSerAspPheGlyLeuAlaLysAla-----GluArgLysGlnLysAspS 383
Db 1292 AGATTGCTGACTTGGCTGCGCCGCGGTCATTGAGAGCAACAGATACAGCGCTGGGAG 1351
Qy 383 eSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLysPheHis 403
Db 1352 GGGGCAAGTCCCATCAAGTGAAGAGCTCTGAAGCATCACTTGGCTCCCTTCACCA 1411
Qy 403 eLysSerAspValTrpSerPheGlyValLeuLeuTrpGluValPheSerTrpArgLys 423
Db 1412 TCAGTGTAGAGCTGTGCTCTTGTGTATCTCTGTATCTGTGATGAGATGTCACCTACGCGGA 1471
Qy 423 lApProTyrrProLysMetSerLeuLysGluValSerGluAlaValGlnLysGlyTyrrArg 443
Db 1472 TCCCTTACCCAGGAGTCAAAACCTGAAAGTATCCGAGTCTGTGAGCGTGTGATACCGGA 1531

```

```

Qy 443 etGluProGluGlnGlyCysProGlyProValHisValLeuMetSerSerCysTrpGlu 463
Db 1532 TGCTCCGCCCAAGAGACTGCCCAAGAGAGCTCTCAACATCATATGCTGCTGGAAAA 1591
Qy 463 lagIuProAlaArgArgProPheArgLysLeuAlaGlnLysLeu-----478
Db 1592 ACCGTCCGAGAGAGCGGCCGCTTCGATTCATTCAGAGTGTGCTGATGATCTTACA 1651
Qy 479 --AlaArgGlnLeuArgSerAla-----GlyAlaProAlaSerValSerg 493
Db 1652 CGGCCACAGAGAGGCACTACCAACAGAGCATGATAGGAGAGGAGGAGGAGGAGG 1711
Qy 493 lGlnAspAlaAspGlySer-----ThrSerProArgSerGlnLysPro 507
Db 1712 GGTCCCAAGTGTGTGCTCGAAGTGGCTCCAGACCATTCGCCAGGAGGCCACACCC 1770

RESULT 9
US-10-007-010-3
: Sequence 3, Application US/10007010
: Publication No. US20030125275A1
: GENERAL INFORMATION:
: APPLICANT: Alexander H. Borchers
: APPLICANT: Kenneth W. Dobie
: TITLE OF INVENTION: ANTISENSE MODULATION OF HCK EXPRESSION
: FILE REFERENCE: RFS-0345
: CURRENT APPLICATION NUMBER: US/10/007,010
: CURRENT FILING DATE: 2001-12-04
: NUMBER OF SEQ ID NOS: 87
: SEQ ID NO 3
: LENGTH: 2015
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (169)...(1686)
US-10-007-010-3

Alignment Scores:
Pred. No.: 3,4e-65 Length: 2015
Score: 725.50 Matches: 180
Percent Similarity: 51.25% Conservative: 87
Best Local Similarity: 34.55% Mismatches: 187
Query Match: 27.16% Indels: 67
DB: 14 Gaps: 15

US-09-977-260-2 (1-507) x US-10-007-010-3 (1-2015)
Qy 34 HisProProValSerAlaArg-----41
Db 272 CATCCACCA-----TCAGCGCGGGCCCTATACCCACACAGCAACACAGCAATCA 325
Qy 42 -----MetProThrArgArgTrpAlaProGlyThrGlnCysIleThrLysCys 57
Db 326 GGGAGCAGGCTGTGAGACATCATCGTGTGCC-----TGT 364
Qy 58 GluHisThrArgPro-LysPro-GlyGlnLeuAlaPheArgLysGlyAspValAlaThrI 77
Db 365 ATGATTTCAGAGCCATTCACCAACAGACCTCACCAAGAGGAGGAGCCAGATGGTGG 424
Qy 77 leuGlnAlaLysGlnAsnLysSerTrpTyrrArgValLysHisHisThrSerGlyGln 97
Db 425 TCCTAGAG-----GAATCCGGGAGTGTGGAGGCTCGATCCCTGGCCACCCGGAAG 478
Qy 97 luglyLeuLeuAlaIaglyAlaLeuArgIuArgIuAlaLeuSerAlaAspProlySL 117
Db 479 AGGGCTACATCCCAAGCAACTATGTCCGCCGTTGACTCTCTGGAGACAG-----531
Qy 117 euserLeuMetProTPrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuG 137
Db 532 -----GAATGGTTTTCAGAGGCAATCAGCCGGAAGAGGAGGAGCCGCACTGC 580
Qy 137 InProProGluAsp-----GlyLeuPheLeuValArgGluSerAlaArgHisProGlyA 155

```

Db	581	TGGCTCCGGGCAACATGCTGGCGCTCCATCATGCTACGATGACGAGACCACTAAAGAA	640
Oy	155	spTyrValLeuLysVal-----SerPheGlyArgAspValIleHisTyrA	170
Db	641	GCTACTCTTTGTCCGTGGCGAGACTACGACCTCGCGAGGGAATACCGGAAACATTAA	700
Oy	170	rgValIleHisArgAsp---GlyHisLeuThrIleAspGluValAlaPhePheCysAsnL	189
Db	701	AGATCCGGACCTGGACAACGGGGGCTTCTACATATCCCGCGAAGCACTTCACGACAC	760
Oy	189	eumEAspMetValAluHisTyrSerLysAspLysGlyAlaIleCysThrLysLeuValA	209
Db	761	TGCAGAGCACTGGTGGACCACTACAAAGGGGAGACAGCGGCTCTGGCAGAAACCTCGG	820
Oy	209	rgProGlyArgLysHisGlyThrLysSerAlaGluGluGluLeuAlaArgAlaGlyTrpL	229
Db	821	TGCCCTGCATG-----TCCTCCAAAGCCCAAGACCTTGGAGAAAGATGCCCTGGG	871
Oy	229	euleAsnHisLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGluGlyIlePheValA	249
Db	872	AGATCCCTCGGGAAATCCCTCAAGCTGGAGAAAGAAATCTGGACCTGGCGAGTTGGGAAG	931
Oy	249	alleGlnGlyLysIleuTr---LeuGlyGlnLysValAlaValLysAsnIleLys---CysA	267
Db	932	TCTGGATGGCCACCTACACAACAGCACACCAGAGTGCGAGTAAGACGATGAAGCCAGGA	991
Oy	267	spValThrAlaGlnAlaPheLeuAspGluThrAlaValaMetThrLysMetGlnHisGluA	287
Db	992	GCATGTCCGATGGAGGCTTCCCTCGGACAGGCCAACGTGTGAATACTCGACGATGACA	1051
Oy	287	snLeuValArgLeuLeuGlyValIleLeuHisGlnGlyLeuTyrIleValIleMetGlnHis	307
Db	1052	AGCTGGCTAACTTCATGCGGGGTCACCAAGAGCCCATCTACATCACTACAGGAGTTCA	1111
Oy	307	alseLysGlyAsnLeuValAsnPheLeuArgThrArgGlyArgAlaLeuValaAsnThrA	327
Db	1112	TGGCCAAAGGAAGCTTGGCTGGACCTTTCGAAAAGTACGAGGGCAGCAAGCAAGCATTCG	1177
Oy	327	lagInLeuLeuGlnPheSerLeuHisValAlaGluGlyMetGluTyrLeuGluSerLysL	347
Db	1172	CAAACTCATTTGACTCTTCACGCCCAAGATGAGAGAGCCATGGCTTCATCGACGAGAGA	1233
Oy	347	ysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSerLysAspLeuValAlaL	367
Db	1232	ACTACATCCACCGAGACCTCCGAGCTGCCAACATCTGTGCTCTGCATCCCTGGGTGTA	1291
Oy	367	ysValSerAspPheGlyLeuAlaLysAla-----GluArgLysGlyLeuAspS	383
Db	1292	AGATTGCTGACTTTGGCTCGCCGCGGCTCATTTGAGACAGACAGATACACGGCTGGGAAG	1351
Oy	383	erSerArgLeuProValLysTrpThrAlaProGluAlaLeuLysHisGlyLysPheItrS	403
Db	1352	GGGCCAAATCCCAATCAAGTGGAGAGCTCTGGAAGCCATCAAACTTGGCTCCCTCACCA	1411
Oy	403	erLysSerAspValTrpSerPheGlyValaLeuLeuTrpGluValaPheSerTyrGlyArgA	423
Db	1412	TCAAGTACGAGCTGTGGTCTTTGGTATCCGCTGATGATGAGATGCTCACTACAGGGCGGA	1477
Oy	423	lapTrpTyrProLysMetSerLeuLysGluValaSerGluAlaValaGluLysGlyTyrArgM	443
Db	1472	TCCCTTACCCAGGGAGTGTCAAAACCCCTGAAGTGAATCCGAGCTCTGAGCGTGAATACCGGA	1531
Oy	443	etGluProProGluGlyLysProGlyProValHisValLeuMetSerSerCysTrpGluA	463
Db	1532	TGCCTCGGCCACGAGAACTGCCACAGAGAGCTCTACACATCATGTATGGCGCTGGCAAA	1591
Oy	463	lagLupProAlaArgArgProPheArgLysLeuAlaGluLysLeu-----	478
Db	1592	ACCGTCCGAGAGAGCGGCCGACCTTCGAATACATCCAGAGTGGTGATGACTCTTACGA	1651
Oy	479	--AlaArgGluLeuArgSerAla-----GlyAlaProAlaSerValSerG	493

QY	1652	CGGCACAGAGAGCCAGTACCAACAGCAGCCATGATGAGGAGACCAAGGAGGAGGAGG	17
Db	1652	CGGCACAGAGAGCCAGTACCAACAGCAGCCATGATGAGGAGACCAAGGAGGAGGAGG	17
QY	493	lyginaspalaasplylser-----ThserProargserGlnGluPro	507
Db	1712	GGTGCCACAGGTGCTGTGGAAGTGTCTCCAGCACCATTCGCCAGAGGCCACACCCCC	1770
RESULT 10			
US-09-917-800A-1611	: Sequence 1611, Application US/09917800A		
US-09-917-800A-1611	: Patent No. US20020119462A1		
GENERAL INFORMATION:			
APPLICANT:	Mendick, Donna		
APPLICANT:	Porter, Mark		
APPLICANT:	Johnson, Kory		
APPLICANT:	Castle, Arthur		
APPLICANT:	Elashoff, Michael		
APPLICANT:	Gene Logic, Inc.		
TITLE OF INVENTION:	Molecular Toxicology Modeling		
FILE REFERENCE:	44921-5038-US		
CURRENT APPLICATION NUMBER:	US/09/917,800A		
CURRENT FILING DATE:	2001-07-31		
PRIOR APPLICATION NUMBER:	US 60/222,040		
PRIOR FILING DATE:	2000-07-31		
PRIOR APPLICATION NUMBER:	US 60/222,880		
PRIOR FILING DATE:	2000-11-02		
PRIOR APPLICATION NUMBER:	US 60/290,029		
PRIOR FILING DATE:	2001-05-11		
PRIOR APPLICATION NUMBER:	US 60/290,645		
PRIOR FILING DATE:	2001-05-15		
PRIOR APPLICATION NUMBER:	US 60/292,336		
PRIOR FILING DATE:	2001-05-22		
PRIOR APPLICATION NUMBER:	US 60/295,798		
PRIOR FILING DATE:	2001-06-06		
PRIOR APPLICATION NUMBER:	US 60/297,457		
PRIOR FILING DATE:	2001-06-13		
PRIOR APPLICATION NUMBER:	US 60/298,884		
PRIOR FILING DATE:	2001-06-19		
PRIOR APPLICATION NUMBER:	US 60/303,459		
PRIOR FILING DATE:	2001-07-09		
NUMBER OF SEQ ID NOS:	1740		
SOFTWARE:	PatentIn Ver. 2.1		
SEQ ID NO 1611			
LENGTH:	1911		
TYPE:	DNA		
ORGANISM:	Rattus norvegicus		
FEATURE:			
OTHER INFORMATION:	Genbank Accession No. US20020119462A1 NM_013185		
US-09-917-800A-1611			
Alignment Scores:			
Pred. No.:	3,43e-64	Length:	1911
Score:	715.50	Matches:	172
Percent Similarity:	54.01%	Conservative:	84
Best Local Similarity:	36.29%	Mismatches:	179
Query Match:	26.79%	Indels:	39
DB:	10	Gaps:	13
US-09-917-260-2 (1-507)	x US-09-917-800A-1611 (1-1911)		
QY	27	ProArgPheLeuAlaArgAlaTrpHisProProAlaSerAlaArgMetProThrArgArg	46
Db	281	CCACGTCGCCCTTAAGAGCTGG---GACGAGACAGCATCAACAGCGTGGCCCGGGGTGG	337
QY	47	Trp---AlaProGlyThrGlnCysIleThrIlyScysGlnHisThrArgPro-LysPro-G	65
Db	338	TGGAGGCTGTGAGACACCATTTGTGTGTCGACTGTACGACTATGAGGCCATTGCACCGTG	397
QY	65	LysIleuAlaPheArgIysGlyAspValValThrIleLeuGluAlaCysGluAsnLys	85
Db	398	AAGACCTCAGCTTCACAGAGGAGACCAAGATGCTGTTCGAG-----GAGTCTGGGG	451
QY	85	eTrpTyArgValLlysHisHisThrSerGlyGlnGluGlyLeuLeuAlaIleGlyAlaL	105


```
Db 452 AGTGGTGAAGGCCGTTCCCTGGCTACCAAGAAAGGCTATATCCCAAGATATATG 511
Qy 105 euaTgGluAArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetProTrpPheHisG 125
Db 512 TAGCTCGAGTTAACTCTTTGGAGACTGAG-----GAGTGGTTCTTCA 553
Qy 125 LyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProGlnLysP-----GlyL 143
Db 554 AGGGATATCAGCCGGAAGAGATGACAGAGCCACCTGCTGGCTCCGGGAACATGCTGGCT 613
Qy 143 euePheLeuValAArgGluSerAlaArgHisProGlyAspTyrValLeuGlyVal----- 160
Db 614 CCTTATCATCTCCGGACAGCTAGACCCAAAGGAGAGCTACTCTTCTGTCAGAGACT 673
Qy 161 -----SerPheGlyAArgAspValIleHisTyrArgValLeuHisArgsp---GlyH 177
Db 674 TTGACCCCGACAGCAGACACGCGTGAAGCATTTAAATCCGACACTGGACAGTGGAG 733
Qy 177 IsLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetValGlnHisTyrS 197
Db 734 GGTCTACATCTCTCCGAGAGACACCTTCACAGCCTGACAGAACTTGCTCCACTCA 793
Qy 197 erLysAspLysGlyAlaIleCysThrLysLeu-----ValArgProLys---A 212
Db 794 AGAAGGGGAAGATGGCTGTGCCAAGAGCTGTCAAGCTCCCTGTGCTCCGAACCC 853
Qy 212 rGlySHisGlyThrLysSerAlaGluGluGluAlaArgAlaGlyTrpLeuLeuAsnL 232
Db 854 AGAAGCCATGGGAGAAAGATGCC-----TGGAGATTTCTC 889
Qy 232 eugInHisLeuThrLeuGlyAlaGlnIleGlyGluGlyGluPheGlyAlaValLeuGlnG 252
Db 890 GAGATCCCTCTCAGATGAGAAAGAACTGGAGCCGCGACATTGGAGAACTGGAGATG 949
Qy 252 LyGluTyr---LeuGlyGlnLysValAlaValLysAsnIleLys---CysAspValTrp 270
Db 950 CCACCTACACAAACACACCAAAAGTGGCGTGAAGACAAATGAACGACAGGAGCATGTG 1009
Qy 270 IagInAlaPheLeuAspGluThrAlaValMetThrLysMetGlnHisGlnAsnLeuVal 290
Db 1010 TGGAGGCTTCTCGGACAGAGGCCAACCTGATGAAGACCTTACAGCATATAACTGGGGA 1069
Qy 290 rGleLeuGlyValIleLeuHisGlnGlyLeuTyrIleValMetGlnHisValSerLysG 310
Db 1070 AGCTAACAGCGTGGTGTCTCAGAGGCCCATCTTATGTGACCGAGTTCATGGCCAAAG 1129
Qy 310 LysAsnLeuValAsnPheLeuArgThrArgIleArgAlaLeuValAsnThrAlaGlnLeu 330
Db 1130 GAAGCCTGCTGCACTTCTCAAGAGTGAAGAGCAGCAACGACCACTGCCAAACTCA 1189
Qy 330 eugInPheSerLeuHisValIleGluGlyMetGluTyrLeuGlnSerLysLysLeuValH 350
Db 1190 TTGACTTCTCAGCCAGATTTCAAGAGGCAATGCGCTTTCATTGACAGAGGAATACATCC 1249
Qy 350 IsArgAspLeuAlaAlaArgAsnIleLeuValSerGluAspLeuValAlaLysValSera 370
Db 1250 ACCGAGACGCTCCGGGCTGCCAACAATCTTGTTCATCCTGCTGTGAAGTCCGCTG 1309
Qy 370 sPrPheGlyLeuAlaLysAla-----GluArgLysGlyLeuAspSerSerArgL 386
Db 1310 ACTTTGGACTGCGACGATCATCGAGACAAATGAGTACACGCTCGGGAAGAGCCCAAGT 1369
Qy 386 eueProValLysTrpThrAlaProGlnAlaLeuLysHisGlyLysPheThrSerLysSera 406
Db 1370 TCCCATATCAAGTGCACAGCTCTTAAGCCATCACTTTGGCTCTTCACTCAAGTAG 1429
Qy 406 sPValTrpSerPheGlyValLeuLeuTrpGluValPheSerTyrGlyArgAlaProTyrP 426
Db 1430 ATGCTGGCTCTTGGATCTGCTGATGAAATCGTCACTACGCGCGGATCCCTTACC 1489
Qy 426 rOlySmetSerLeuLysGluValSerGluAlaValGluLysGlyTyrArgMetGluProP 446
Db 1490 CAGGTATGTCAAACCCAGAGGTGATTCGAGCACTAGACATGAGGTACCGTATGCCCTGCAC 1549
```

```
Qy 446 roGluGlyCysProGlyProValHisValLeuMetSerCysTrpGluAlaGluProA 466
Db 1550 CAGATTAAGTCCCGCAGAGAGAGCTTACACTATATCATGATCCGCTGGAGAAACGCTCAG 1609
Qy 466 IaArgArgProProPheArgLysLeuAlaGluLysLeu 478
Db 1610 AGGAAGGCGCCACTTTCGATATACATCCAGACGCTGCTG 1647

RESULT 11
US-10-298-377A-3
; Sequence 3, Application US/10298377A
; Publication No. US20030130209A1
GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Cheresn, David A.
; APPLICANT: Paul, Robert
; APPLICANT: Elicelri, Brian
; TITLE OF INVENTION: Method of Treatment of Myocardial
; FILE REFERENCE: TSRI-651.5
CURRENT APPLICATION NUMBER: US/10/298,377A
CURRENT FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: 10/298,377
PRIOR FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: 09/470,881
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 09/538,248
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: PCT/US99/11780
PRIOR FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: 60/087,220
PRIOR FILING DATE: 1998-05-29
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 4517
TYPE: DNA
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (208)...(1839)
US-10-298-377A-3

Alignment Scores:
Pred. No.: 4,21e-63 Length: 4517
Score: 710.00 Matches: 163
Percent Similarity: 54.61% Conservative: 86
Best Local Similarity: 35.75% Mismatches: 165
Query Match: 26.58% Indels: 42
DB: 15 Gaps: 11

US-09-977-260-2 (1-507) x US-10-298-377A-3 (1-4517)
Qy 66 GluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAlaCysGluAsnLysSer 85
Db 532 GACCTTCATTTAAGAAAGGTGAAGATTTCAAAATTAATTAACATACATGGAAGGA---GAT 588
Qy 86 TrpTyrArgValLysHisHisThrSerGlyGlnGlyLeuLeuAlaIleGlyAlaLeu 105
Db 589 TGTGGGAAGCAAGATCAATGCTAACAGAAAGATGTTATTCGCGCAAAATTAATGA 648
Qy 106 ArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetProTrpPheHisGly 125
Db 649 GCGCGTCGAGATTCATTCACGCAAGAA-----GAATGGTATTTTGGC 690
Qy 126 LysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProPro-----GluAspGlyLeu 143
Db 691 AAATGGGGAAGAAAGATGCTGAAGATTAATCTTTGAATCCTCGGAATCAACAGATATT 750
Qy 144 PheLeuValAArgGluSerAlaArgHisProGlyAspTyrValLeuCysValSer----- 161
Db 751 TTCTTAGTAAGAGAGAGTGAAGCAACTAAAGTGTCTTATTCCTTGTATTCGTGATTTGG 810
```

QY 162 -----PheGlyArgAspValIleHisThrArgValLeuHisArgAsp---GlyHis 177
 Db 811 GAGGAGATTAAGGGGTGACAAATGTAAACACTACAAATTTGGAACACTTGCAATGCTGGA 870
 QY 178 LeuThrIleAspGluAlaValIlePhePheCysAsnLeuMetAsnValGluHisTyrSer 197
 Db 871 TACTATATCAACACACAGACCAATTTGATCTGTGCAAGAAATTTGGTGAACACTACACA 930
 QY 198 LysAspLysGluAlaIleCysThrLysLeu-----ValArgProLysArgLysHisGly 215
 Db 931 GAACATGCTGATGTATGCTCCACAAAGTGTGATGCAAACTGTGTCCAACTGTGAACCTGAG 990
 QY 216 ThrLysSerAlaGluGluGluLeuAlaArgAlaGlyTyrProLeuAsnLeuGluHisLeu 235
 Db 991 ACTCAAGT-----CTAGCAAAAGATGCTTGGGAATCCCTCGAGATCTTGG 1038
 QY 236 ThrLeuGlyAlaGluIleGlyGluGlyIlePheGlyAlaValIleGluGluGlyTyrLeu 255
 Db 1039 CGACTAGAGGTTAACTAGACAAAGATGTTTGGCCCAAGTGTGATGAGGAGACATGGAAT 1098
 QY 256 Gly----GlnLysValAlaValLysAsnIleLys---CysAspValThrAlaGluAlaPhe 273
 Db 1099 GGAACSCAGAAAGTACGCAATCAAAACACATCAACAGTACATGATGCCAGAAAGCTTTC 1158
 QY 274 LeuAspGluThrAlaValIleMetThrLysMetGlnHisGluAsnLeuValArgLeuLeuGly 293
 Db 1159 CTTCAGAAAGCTCAGATATGAAATAATTAAGACATGATTAACCTGTCTCCATATATGCT 1218
 QY 294 ValIleLeuHisGlnGlyLeuTyrIleValIleMetGlnHisValSerLysGluAsnLeuVal 313
 Db 1219 GTTGTCTTCAGAAACCAATTTACATTTGCTACAGTAAATTTATGTCMAAGAGACCTTATTA 1278
 QY 314 AsnPheLeuArgThrArgGlyArgAlaLeuValAsnThrAlaGluIleLeuGluIlePheSer 333
 Db 1279 GATTTCTTAAAGAGAGAGATGCAAAATGATTTGCAAGCTTCCACAGCTGTGATATGGCT 1338
 QY 334 LeuHisValAlaGluGlyMetGluTyrLeuGluSerLysLysLeuValHisArgAspLeu 353
 Db 1339 GCTCAAGTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1398
 QY 354 AlaAlaArgAsnIleLeuValSerGluAspLeuValAlaValSerAspPheGlyLeu 373
 Db 1399 CGGGCTGCTATATCTTGTAGAGAAATCTGTGTGCAAAATACAGACCTTGTGTTTA 1458
 QY 374 AlaLysAla-----GluArgLysGlyLeuAspSerArgLeuProValLys 389
 Db 1459 GCAAGGTTAATTGACAAATGAAATACACAGCAAGCAAGCTGCAAAATTTCCAAATCAAA 1518
 QY 390 TrpThrAlaProGluAlaLeuLysHisGlyLysPheThrSerLysSerAspValTrpSer 409
 Db 1519 TGGACAGCTCCTGAAGCTGACCTGTATGCTGCTGTACAAATTAAGTGTGATGCTGTGCTCA 1578
 QY 410 PheGlyValLeuLeuLeuTrpGluValIlePheSerTyrGlyArgAlaProTyrProLysMetSer 429
 Db 1579 TTTGGAATTCCTGCAACACAGACTAGTAACAAGAGGCGGAGTGCATATCCAGTATAGGTC 1638
 QY 430 LeuLysGluValSerGluAlaValGluLysGlyTyrArgMetGluProProGluGlyCys 449
 Db 1639 AACCGTGAAGTACTAACAAGTGGAGCGAGGTATACAGATGCCCTGCGAGGCTGT 1698
 QY 450 ProGlyProValHisValLeuMetSerSerCysTrpGluAlaGluProAlaArgPro 469
 Db 1699 CCAAGAACTCCCTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1758
 QY 470 ProPheArgLysLeuAlaGluLysLeuAlaArgGluLeuArgSerAlaGlyAlaProAla 489
 Db 1759 ACATTTGATATATTCAGCTCTTCTTG----- 1785
 QY 490 SerValSerGlyGlnAspAlaAspGlySerThrSerProArgSerGln 505
 Db 1786 -----GAAGACTACTTCACTGCTACAGGCCACAGTACCAG 1821

RESULT 12
 US-10-175-523-50
 ? Sequence 50, Application US/10175523
 ? Publication No. US20030096264A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Brockman, Jeffrey
 ? APPLICANT: Evans, David
 ? APPLICANT: Hook, Derek
 ? APPLICANT: Klimczak, Leszek
 ? APPLICANT: Laeng, Pascal
 ? APPLICANT: Palfreyman, Michael
 ? APPLICANT: Rajan, Prithi
 ? TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPTS)
 ? FILE REFERENCE: 3235/10795-US3
 ? CURRENT APPLICATION NUMBER: US/10/175,523
 ? PRIOR FILING DATE: 2002-06-18
 ? PRIOR APPLICATION NUMBER: US 60/299,151
 ? PRIOR FILING DATE: 2001-06-18
 ? PRIOR APPLICATION NUMBER: US 60/317,828
 ? PRIOR FILING DATE: 2001-09-07
 ? PRIOR APPLICATION NUMBER: US 60/325,150
 ? PRIOR FILING DATE: 2001-09-25
 ? PRIOR APPLICATION NUMBER: US 60/333,047
 ? PRIOR FILING DATE: 2001-11-14
 ? PRIOR APPLICATION NUMBER: US 60/349,936
 ? PRIOR FILING DATE: 2002-01-18
 ? PRIOR APPLICATION NUMBER: US 60/361,834
 ? NUMBER OF SEQ ID NOS: 197
 ? SOFTWARE: PatentIn version 3.1
 ? SEQ ID NO: 50
 ? LENGTH: 2298
 ? TYPE: DNA
 ? ORGANISM: Homo sapiens
 ? US-10-175-523-50
 Alignment Scores:
 ? Pred. No.: 1,85e-63 Length: 2298
 ? Score: 709.50 Matches: 167
 ? Percent Similarity: 53.23% Conservative: 97
 ? Best Local Similarity: 33.67% Mismatches: 179
 ? Query Match: 26.56% Indels: 53
 ? DB: 14 Gaps: 14
 US-09-977-260-2 (1-507) x US-10-175-523-50 (1-2298)
 QY 35 ProProValSerAlaArgMetProThrArgArgTrp-----Ala 48
 Db 427 CGAGTCCAGAAATCTCAGCTTTTACCTGGACAGAGTTTCAAACTAAAGTCCAGAGAA 486
 QY 49 ProGlyThrGlnCysIleThrLysCysGluHisThrArgProLysProGlyLeuAla 68
 Db 487 CAAGGACATTTGTGATCCCTTGTACCCCTATGATGATGATCCACCCGAGCACTGTCT 546
 QY 69 PheArgLysGlyAspValValThrIleLeuGluAlaCysGluAsnLysSerTrpArg 88
 Db 547 TTCAGAAAGAGAGAGATGAAAGTCTGGAG-----GACATGAGCAATGCTGGAAA 600
 QY 89 ValLysHisThrSerGlyGluGluGluLeuAlaGlyAlaLeuArgGluArg 108
 Db 601 GCAAACTCCCTTTTAAACAAAAAGAGGCTTCATCCCAAGAACTATGTGGCCAAATCTC 660
 QY 109 GluAlaLeuSerAlaAspProLysLeuSerLeuMetProTrpPheHisGlyLysIleSer 128
 Db 661 AACACCTTGAAGAAAGAA-----GACTGGTTTTTCAAGATATTAAC 702
 QY 129 GlyGlnGluAlaValGlnLeuGluIleProProGluAsp-----GlyLeuPheLeuVal 146
 Db 703 AGCAAGAGAGCAGAAAGCAGCTTTTGGACACAGAAATAGCGGTGAGACTTTCTTAT 762
 QY 147 ArgGluSerAlaArgHisProGlyAspArgValIleCysValSer----- 161
 Db 763 AGAGAAAGTGAACATTTAAAGAGCTTCTCTGTCTGTGACAGACTTTGACCCCTGTG 822

QY 154 GLYASPTyValLeucyValSerPheGlyArgAspValIleHisTyArgVal---Leu 172
||| : : : : :
Db 622 GGCCAGAGGTCCATCTCGGTGAGATGAGAGGAGGAGGTGATACAGACACACT 681
QY HisArgAspGlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMet 192
||| : : : : :
Db 682 GCTTCGATGCGCAGAGCTTACGTCCTCTCCGAGACGCCGTTCAACACCCCTGGCCGAGTTG 741
QY ValGlnHisTySerLeysAspLysGlyAlaIleCysThrLysLeu-----ValArgPro 210
||| : : : : :
Db 742 GTTCATCATCATTCATCAACGCTGCGCCGAGGCTCATGACCCAGCTCATATTCAGGCCCA 801
QY 211 LysArgLys-----HisGlyThrLysSerAlaGluIleGluLeuAlaArgAla 226
||| : : : : :
Db 802 AACGGGACACAGCCCACTGTATGAGTGTCTCCCACTACACAG----- 849
QY 227 GATPLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGluIlePhe 246
||| : : : : :
Db 850 ---TGGGAGATGAGACCGCAGCATCACCATGACACACAGCTGGCGGGGCGCAGTAC 906
QY 247 GlyAlaValLeuGlnGlyLysLys-----LeuGlyGlnLysValAlaValLysAsnIle 264
||| : : : : :
Db 907 GGGGAGGTGTACAGCGGCGCTGTGAGAAATACAGCCTGACGCTGGCGCTGAGAGACTTG 966
QY 265 LysCysAsp---ValThrAlaGlnAlaPheLeuAspGluThrAlaValMetThrLysMet 283
||| : : : : :
Db 967 AAGGAGGACACCCATGAGAGGTGAGAGTCTTGAAGAACCTGACATGAGAAAGAGATC 1026
QY 284 GlnHisGlnAsnLeuValArgLeuLeuGlyValIleLeuHisGln---GlyLeuTyIle 302
||| : : : : :
Db 1027 AACACCCCTTACCTGGTGCAGCTCTTGGGCTGTGACCCGGAGGCCGCTTCTATATC 1086
QY 303 ValMetGlnHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgAla 322
||| : : : : :
Db 1087 ATCAGTAGTCTACCTAGCTACAGGAGACCTCTGACTGAGGAGGTGACACCGGAG 1146
QY 323 LeuValAsnThrAlaGlnLeuGlnPheSerLeuHisValAlaGluIleMetGlyLys 342
||| : : : : :
Db 1147 GAGGTACAGCCCGTGTCTGTCTGATACAGGCGCAGATCTCGTACAGCCATGAGATAC 1206
QY 343 LeuGlnSerLysLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSerGlu 362
||| : : : : :
Db 1207 CTGGAGAGAGAAATCTTCATCCACAGATCTTCTGCGCCGAAACCTGCTGATGAGGAG 1266
QY 363 AspLeuValAlaLysValSerAspPheGlyLeuAlaLys----- 375
||| : : : : :
Db 1267 AACCACTGTGTGAGAGTACGTGATTTGGCTGAGCAGGTGATGACAGGGGACACTTAC 1326
QY 376 AlaGluArgLysGlyLeuAspSerSerArgLeuProValLysThrPheAlaProGluAla 395
||| : : : : :
Db 1327 AACGCCCATGCTGA-----GCCAAGTCCCATCAATGAGTGCACCCGAGAGC 1377
QY 396 LeuLysHisGlyLysPheThrSerLysSerAspValThrSerPheGlyValLeuLeuTrp 415
||| : : : : :
Db 1378 CTGGGCTTACACAGATCTTCATCAAGTCCGAGCTGTGGCATTTGGAGATTTGCTTGG 1437
QY 416 GluValPheSerTyGlyArgAlaProTyProLysMetSerLeuLysGluValSerGlu 435
||| : : : : :
Db 1438 GAATTTGTAACCTATGCGATGTCCCTTACCCGCGAATGATGACTGTCCAGAGTGTATGAG 1497
QY 436 AlaValGluLysGlyArgMetGluProProGluIleCysProGlyProValHisVal 455
||| : : : : :
Db 1498 CTGCTAGAGAGAGACTTACCGCATGAGCGCCAGAGGCTGCGCCAGAGAGACTTATGAA 1557
QY 456 LeuMetSerSerCysTrpGluAlaGluProAlaArgArgProProPheArgLys----- 473
||| : : : : :
Db 1558 CTTCATCGAGAGCATGTTGGCAGTGGGAATCCCTCTGACCGGCGCTCTTGTCTGAATCCAC 1617
QY 474 -----LeuAlaGluLysLeuAlaArgGluLeu 482
||| : : : : :
Db 1618 CAAGCTTTGAAACAATGTTCCAGAAATCAGATATCTCAGAGCAAGTGGAAAGAGAGCTG 1677
QY 483 ArgSerAlaGlyAlaProAlaSerValSer 492

Db 1678 GGGAAACAAAGCGCTCGCTGGGCGCTGAGT 1707
RESULT 14
US-09-967-768A-300
; Sequence 300, Application US/09967768A
; Patent No. US2002015087A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Slight
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 300
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-300
Alignment Scores:
Pred. No.: 1 52e-60 Length: 2354
Score: 681.50 Matches: 176
Percent Similarity: 49.03% Conserved: 78
Best Local Similarity: 33.98% Mismatches: 198
Query Match: 25.51% Indels: 67
Gaps: 13
US-09-977-260-2 (1-507) x US-09-967-768A-300 (1-2354)
QY 23 ProArgValSerProArgPheLeuArgAlaTrp----- 33
||| : : : : :
Db 185 CCACGG-----CCAAAGAGAGATGTGCGCTGGAGAGGAGACTTCAGAGCTACGGGAG 238
QY 34 -----HisPro-----Pro 36
||| : : : : :
Db 239 CAGACCACTATGCGCTGACCCCACTAAGCCCGCGCTGATCTCATTTGCCACATCC 298
||| : : : : :
QY 37 ProValSerAlaArgMetProThrArgArgTrpAlaProGly----- 50
||| : : : : :
Db 299 CCAACTACACCACTTCTCTCTCAGGCCATC-AACCTGTGCTTCTGTATAGTGGCACC 357
||| : : : : :
QY 51 -----ThrGlnCysIleThrLysCysGlnHisThrArg 61
||| : : : : :
Db 358 ATCAGGAGGTGTGACGAGATTTGGGTGAGCCCTGTCTCATCTGCTGATGAGGCT 417
||| : : : : :
QY 62 ProLysProGluIleLeuAlaPheAlaGlyGlyAspValValThrIleLeuGluAlaCys 81
||| : : : : :
Db 418 CGAACTGAGAGATCTCACTCACTTCCACCAAGGGGAGAAAGTTCCACATCTGGAACAATACT 477
||| : : : : :
QY 82 GluAsnLysSerTyArgTrpValLysHisIleThrSerGlyGlnGluLeuAla 101
||| : : : : :
Db 478 GAAAGT---GACTGGTGGAGAGCTCGCTCTCAGCTCCGGAAATCTGGTCTTCCAGCGCAAC 534
||| : : : : :
QY 102 AlaGlyAlaLeuArgGluArgGluAlaLeuSerAlaAspProLysLeuSerLeuMetPro 121
||| : : : : :
Db 535 AGCAACTAGGTGGCCCTGTGTACATCAATCAAGTGA-----GAG 576
||| : : : : :
QY 122 TrpPheHisGlyLysIleSerGlyGlnGluAlaValGlnGlnLeuGlnProPro----- 139
||| : : : : :
Db 577 TGGTACTTTGGAAAGATTGGAGAAAGATGACAGAGGACAGCTGCTTCCACCGGCAAC 636
||| : : : : :
QY 140 GluAspGlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyValLeuCys 159
||| : : : : :
Db 637 CCCAGGGGCGCTTCTCATTCGAGAAAGCGAGACCCCAAGAGTCTTACTCTCTGCTCC 696

QY 160 Val-----SerPheGlyArgAspValIleHisTyrArgValIleuHisArg 174
Db 697 ATCCGGAGCTGGGATCAGACAGGCGGATCATGTGAAGCATTCAGAAAGTCCGAAATCTG 756
QY 175 Asp---GlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetVal 193
Db 757 GACATGGCGGCTACTACATCACCACAGCGGATTCAGTTCAACAGTGGTGACAGGAGCTGTG 816
QY 194 GluHisTyrSerIleAspIleGlyAlaIleCysThrIleuValAlaProIleArgLys 213
Db 817 CAGACATACATGAGGTGAATGACGGCTGTGCAACCTGTCTCAACGCGCTCCATCAGCATTG 876
QY 214 HisGlyThrIleSerIleGluIleGluAlaValIleGluAlaValIleGluAlaValIleGlu 233
Db 877 -----ATGAAGCCGCGAGAGCTGGCTGGCCCAAGAGCGCTGGAGATCAGCCGAGC 930
QY 234 HisLeuThrIleuGlyAlaGlnIleGlyGluIleGluIleGluAlaValIleGluIleGlu 253
Db 931 TCCATCAGCTGGAGAGCGCGCTGGCAGCGGCTGTCTGGGAGTGTGTGGCTGGGAGC 990
QY 254 TyrLeuGlyGln---LysValAlaValIleAsnIleLys---CysAspValIleAlaGln 271
Db 991 TGGAAACGCGAGCACTAAGGTGGCGGTGAAGCGCTGAAGCGCGGACACCATGTCCCGAG 1050
QY 272 AlaPheLeuAspGluThrAlaValIleMetThrIleuMetGlnHisGluAsnLeuValArgLeu 291
Db 1051 GCCTCTCTGGAGAGAGCGCGAGCTATGACGTGTGGCGCAAGCAAGCTGTGTGCTGCTG 1110
QY 292 LeuGlyValIleLeuHisGlnIleGlyLysIleuValIleMetGlnHisValIleSerLysGlyAsn 311
Db 1111 TACCGCTGGTGTGGAGAGCGCTGATCTACATCTGAGCGATGTGATGTGTCAGCGGAGC 1170
QY 312 LeuValAsnPheLeuAlaGlyThrArgGlyAlaArgAlaLeuValAsnThrAlaGlnLeuGln 331
Db 1171 TTCTGGATTTTTCAGAAACCCAGAGCGCGAGATTTGAGCGTCCGCCCAATGTGTGAC 1230
QY 332 PheSerLeuHisValAlaGluIleGlyMetGlyLysLeuIleuSerLysLysValIleHisArg 351
Db 1231 ATGGCAGCCAGAGTACGTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290
QY 352 AspLeuAlaAlaArgAsnIleLeuValIleSerLysLeuValAlaValIleValSerAspPhe 371
Db 1291 GACCTGAGCGCAGCAACATCTGTGTGGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1350
QY 372 GlyLeuAlaLysAla-----GluArgLysGlyLeuAspSerSerArgLeuPro 387
Db 1351 GCGTGGCGGCTCTCATCAAGAGAGATGATGACAAACCCCTGCCAAGTTCCCAAGTTCCCC 1410
QY 388 ValIleTyrThrAlaProGluAlaLeuLysHisGlyLysPheThrSerIleSerAspVal 407
Db 1411 ATCAAGTGCAGACGCCCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1470
QY 408 TrpSerPheGlyValIleLeuLeuTyrGluValPheSerTyrGlyAlaGlnProTyrProLys 427
Db 1471 TGGCTCTTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1530
QY 428 MetSerLeuLysGlyValIleSerGluAlaValIleGluLysGlyTyrArgMetGluProGlu 447
Db 1531 ATGATATAAAGGAGAGTGTGGAACAGAGTGAAGAGGAGGAGTGAACAGAGTGTGAGCCGCTCA 1590
QY 448 GlyCysProGlyProValHisValLeuMetSerSerSerSerSerSerSerSerSerSer 1650
Db 1591 GGGCGCCAGAGTCCCTGTCAGAGGAGCATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1650
QY 468 ArgProPheArgLysLeuAlaGluLysLeuAlaArgLysLeuAlaArgLysLeuAlaArg 485
Db 1651 AGGCGTACCTGAGTACCTGAGTCTCTGAGAGAGTACTTCACTCCGCT 1704

RESULT 15
US-09-977-269-5
; Sequence 5, Application US/09977269
; Patent No. US20020082037A1

GENERAL INFORMATION:
; APPLICANT: ULRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977, 269
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232, 545
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2770
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (366)..(1880)
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; US-09-977-269-5
Alignment Scores:
Pred. No.: 1,2e-57 Length: 2770
Score: 654.50 Matches: 153
Percent Similarity: 53.068 Conservative: 81
Best Local Similarity: 34.698 Mismatches: 166
Query Match: 24.508 Indels: 41
DB: 9 Gaps: 12
US-09-977-260-2 (1-507) x US-09-977-269-5 (1-2770)
QY 66 GluLeuAlaPheArgLysGlyAspValValThrIleLeuGluAlaCysGluAsnLysSer 85
Db 543 GACTTGGAGTTCAGAGCAGGAGGAGCAACATTCAGATCTGAGCACTTGG-----CATGAGCGC 599
QY 86 TrpIleArgValLysHis-----HisThrSerGlyGlnGly 98
Db 600 TGGTGGTTGGCCAGACACTTGGAGAAAGAGAGATGGCTCCAGTCAAGCAATCAAGGC 659
QY 99 LeuLeuAlaAlaGluAlaLeuArgGluAlaLeuSerAlaAspProLysLeuSer 118
Db 660 TATATTCCTCTACATCACTGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 707
QY 119 LeuMetProTyrPheHisGlyLysIleSerGlyGlnAlaValGlnIleGlnPro 138
Db 708 -----CCGTGGTCTTGGAGCAATCGAAGATCAGATCGAGAAACAACTATTTAT 761
QY 139 ProGluAsp-----GlyLeuPheLeuValArgGluSerAlaArgHisProGlyAspTyr 156
Db 762 TCAGAAAACAGACCGGCTTCTCTTAATAGAGAAAGTGAAGAAAGCAAGAAAGAGAAATTC 821
QY 157 ValLeuCysValSerPheGlyArgAspValIleHisTyrArgValLeuHisArgAsp--- 175
Db 822 TCTCTTCACTTAAAGTGAAG 881
QY 176 GlyHisLeuThrIleAspGluAlaValPhePheCysAsnLeuMetAspMetValGluHis 195
Db 882 GGGGAGATTTTTCACAGCGAGAGAAAGATCTTCAACACTGAAAGCAAGATTTGTGAGCCAC 941
QY 196 TyrSerIleAspLysGlyAlaIleCysThrIleuValAlaArgProLysArgLys----- 213
Db 942 TACACCAAGACAACTGACGCGCTGTGTGTGTCAGAGTGGAGAAACATCTTAAAGATTCAG 1001
QY 214 -----HisGlyThrLysSerAlaGluIleGluAlaValArgAlaGly 227
Db 1002 GTCCAGCTGCATTTGATTTGTGCTATATAAAGCTGAGCA----- 1043
QY 228 TrpLeuLeuAsnLeuGlnHisLeuThrLeuGlyAlaGlnIleGlyGluIleGluPheGly 247
Db 1044 TGGAGATAGACCGCAACTCCATACAGCTTCTGAAAGCAAGATTTGGATCTGGTCACTTGGC 1103

```
QY 248 AlaValLeuGlnGlyIuTYrLeuGly---GlnLysValAlaValLysAsnIleLys--- 265
Db 1104 GAAGTATGGGAGAGTGTGTGGAAACAAATACCACATCCAGTACAGTGAACATTAACCA 1163
QY 266 CysAspValThrAlaIleAspLeuAspLeuThrAlaValMetThrLysMetGlnHis 285
Db 1164 GGTTCATGGATCCAAATGACTTCTGAGGAGGACACAGATAAATGAAGAACTTAAGCAT 1223
QY 286 GluAsnLeuValArgLeuGlyVal---IleLeuHisGlnGlyLeuTYrIleValMet 304
Db 1224 CCAAGGTTATCCAGCTTTATGCTTTGCTTTGACCTTTGAGACATCCAAATTTATATTACA 1283
QY 305 GluHisValSerLysGlyAsnLeuValAsnPheLeuArgThrArgGlyArgAlaLeuVal 324
Db 1284 GAGTTGATGACATGGAAGTCTGCAAGAATATCTCCAAATATGACACATGATCAAAAATC 1343
QY 325 AsnThrAlaGlnLeuLeuGlnPheSerLeuHisValAlaGlnGlyMetGlyTYrLeuGlu 344
Db 1344 CATCTAGCTCAACAGGTAGACATGGCGGCACAGGTTGCCCTGGAATGGCTATCTGGAG 1403
QY 345 SerLysLeuValHisArgAspLeuAlaAlaArgAsnIleLeuValSerGluAspLeu 364
Db 1404 TCTCGGAACTACATTCACACAGATCTGGCTGCCAGAAATGCTCCTGTTGGTGAACATAT 1463
QY 365 ValAlaLysValSerAspPheGlyLeuAla-----LysAlaGluArgLysGlyLeu 381
Db 1464 ATCTACAAAGTACAGATTTTGAGCTGCCAGAGTTTTPAAGTAGATATGAAGACATC 1523
QY 382 AspSerSerArg-----LeuProValLysThrPheThrAlaProGluAlaLeuLys 397
Db 1524 TATGAATCTAGACACGAAATAAAGCTGCCGGTAGAGTGAAGTGCCTGCCGAAAGCCATTCGT 1583
QY 398 HisGlyLysPheThrSerLysSerAspValTrpSerPheGlyValLeuLeuTrpGluVal 417
Db 1584 AGTAAATAAATTTCAGCATTAAGTCCGATGTATGCTCATTTGGATCCTCTTTATGAATC 1643
QY 418 PheSerTYrGlyArgAlaProTYrProLysMetSerLeuLysGluValSerGluAlaVal 437
Db 1644 ATTACTTATGGCAAAATGCTTACAGTGTATGACAGTGCCAGGTAATCCAGATGTG 1703
QY 438 GluLysGlyTYrArgMetGluProProGluGlyCysProGlyProValHisValLeuMet 457
Db 1704 GCTCAAACTATAGACTTCGGCAACCATCCACTGACACAGCAATTTTACAACATCATG 1763
QY 458 SerSerCysTrpGluAlaGluProAlaArgArgProPheArgLysLeuAlaGluLys 477
Db 1764 TTGGAGTGGTGAATGCAGAGCCTTAAGGAACGACTACATTTGAGACACTGCTGGAAA 1823
QY 478 Leu 478
Db 1824 CTT 1826
```

Search completed: August 2, 2003, 05:12:26
Job time : 1007 secs